

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 01:36:54 ; Search time 175 Seconds  
(without alignments)

2651.561 Million cell updates/sec

Title: US-10-619-992-2

Perfect score: 6215

Sequence: MPVQAAQWTEFLSSPCDLYNE.....MSEDKNDFLKPVANGRMVNS 1191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 399608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AI:  
 1: /cggn2\_6/ptodata/2/pubpaas/US07\_PUBCOMB.pep.\*  
 2: /cggn2\_6/ptodata/2/pubpaas/PCTN\_NEWPUB.pep.\*  
 3: /cggn2\_6/ptodata/2/pubpaas/US06\_NEWPUB.pep.\*  
 4: /cggn2\_6/ptodata/2/pubpaas/US06\_PUBCOMB.pep.\*  
 5: /cggn2\_6/ptodata/2/pubpaas/PCTUS\_NEWPUB.pep.\*  
 6: /cggn2\_6/ptodata/2/pubpaas/PUBCOMB.pep.\*  
 7: /cggn2\_6/ptodata/2/pubpaas/US08\_NEWPUB.pep.\*  
 8: /cggn2\_6/ptodata/2/pubpaas/US09A\_PUBCOMB.pep.\*  
 9: /cggn2\_6/ptodata/2/pubpaas/US09B\_PUBCOMB.pep.\*  
 10: /cggn2\_6/ptodata/2/pubpaas/US09C\_PUBCOMB.pep.\*  
 11: /cggn2\_6/ptodata/2/pubpaas/US09\_NEWPUB.pep.\*  
 12: /cggn2\_6/ptodata/2/pubpaas/US10A\_PUBCOMB.pep.\*  
 13: /cggn2\_6/ptodata/2/pubpaas/US10B\_PUBCOMB.pep.\*  
 14: /cggn2\_6/ptodata/2/pubpaas/US10C\_PUBCOMB.pep.\*  
 15: /cggn2\_6/ptodata/2/pubpaas/US10D\_PUBCOMB.pep.\*  
 16: /cggn2\_6/ptodata/2/pubpaas/US10E\_PUBCOMB.pep.\*  
 17: /cggn2\_6/ptodata/2/pubpaas/US11A\_PUBCOMB.pep.\*  
 18: /cggn2\_6/ptodata/2/pubpaas/US11B\_PUBCOMB.pep.\*  
 19: /cggn2\_6/ptodata/2/pubpaas/US11C\_PUBCOMB.pep.\*  
 20: /cggn2\_6/ptodata/2/pubpaas/US60\_NEWPUB.pep.\*  
 21: /cggn2\_6/ptodata/2/pubpaas/US60\_NEWPUB.pep.\*  
 22: /cggn2\_6/ptodata/2/pubpaas/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6215	100.0	1191	9	US-09-921-099-2
2	6207	99.9	1191	9	US-09-921-099-2
3	611	9.8	522	15	US-09-921-099-4
4	490	7.9	95	9	US-10-104-047-3665
5	311	5.0	67	9	US-09-864-761-48290
6	203	3.3	2377	18	US-10-840-512-196
7	203	3.3	1267	15	US-10-092-900A-12B
8	200.5	3.2	728	17	US-10-732-923-6679
9	198.5	3.2	3586	16	US-10-334-143-77
10	196.5	3.2	1248	16	US-10-437-963-126940
11	195.5	3.1	123	9	US-09-864-761-42017

SEQUENCES

```

Sequence 9, Appli
Sequence 9, Appli
Sequence 2, Appli
Sequence 40, Appli
Sequence 230, App
Sequence 1404, App
Sequence 140, App
Sequence 224, App
Sequence 145, App
Sequence 119, App
Sequence 4, Appli
Sequence 3, Appli
Sequence 1940, App
Sequence 125036, App
Sequence 3473, App
Sequence 1085, App
Sequence 2, Appli
Sequence 425, App
Sequence 425, App
Sequence 62, Appli
Sequence 589, App
Sequence 7758, App
Sequence 156135, App
Sequence 104121, App
Sequence 140753, App
Sequence 160683, App
Sequence 14, Appli
Sequence 14, Appli
Sequence 2228, App
Sequence 32198, A
Sequence 181703, A
Sequence 188268, A
Sequence 195106, A
Sequence 2246, App
US-09-921-099-2
RESULT 1
US-09-921-099-2
; Sequence 2, Application US/09321099
; Patent No. US20020107372A1
; GENERAL INFORMATION:
; APPLICANT: Hefneider, Steven
; APPLICANT: Markins, Louise
; APPLICANT: Bennett, Donald
; APPLICANT: Seiss, Donald
; TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene
; FILE REFERENCE: 00-617-A
; CURRENT APPLICATION NUMBER: US/09/921,099
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 2
; LENGTH: 1191
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-921-099-2

```

ALIGMENTS

Query	Match	Score	Best Local Similarity	Matches	Pred. No.	Mismatches	Indels	Gaps	0;
Qy	1	MPVQAAQNTTEFLSPCPTYNEFEDVNHKPISLGCSTHTVKTCNLKHLRACPPFDGTAINTD	60						
Db	1	MPVQAAQNTTEFLSPCPTYNEFEDVNHKPISLGCSTHTVKTCNLKHLRACPPFDGTAINTD	60						
Qy	61	IDVPVNFAFLQLVGAQVDPHQSTKLSNLGENKHYEVAKKCVDIALYKPLSGKGVAS	120						
Db	61	IDVPVNFAFLQLVGAQVDPHQSTKLSNLGENKHYEVAKKCVDIALYKPLSGKGVAS	120						
Qy	121	LNGQSLSRPMQRKLTVLYNCQLVEEEGRVTRAMRAARSIGERTVTBLILHQNPQLSANL	180						

Page 2

1.21	LNQSSLRPWRKLTLLVNCOLVEEGVRAMRAARSGLGERTVTVELQHQNPQLSANL	180
1.81	WAVERARGCOPLGPMQBEALKLVIALEDLSASRKLVLFLVFFVQBLPREPOQSSTSIG	240
1.81	WAVERARGCOPLGPMQBEALKLVIALEDLSASRKLVLFLVFFVQBLPREPOQSSTSIG	240
2.41	HVQOLLYRASCFCFKYTKRDEDSMOLKEEPRSYEARREHDAQTIVHAMEAGLRLISPEQW	300
2.41	HVQOLLYRASCFCFKYTKRDEDSMOLKEEPRSYEARREHDAQTIVHAMEAGLRLISPEQW	300
3.01	SSLJYGDLAHKSHMQS1IDKQSPBSRAKSQELTIVLQRTGDPANLNRPHFLJELANI	360
3.01	SSLJYGDLAHKSHMQS1IDKQSPBSRAKSQELTIVLQRTGDPANLNRPHFLJELANI	360
3.61	DPNPDAVSPTEQLENAMAVKTVHGLVDFPIQNSRKGHETPQPOPNSKXTSMGRDLR	420
3.61	DPNPDAVSPTEQLENAMAVKTVHGLVDFPIQNSRKGHETPQPOPNSKXTSMGRDLR	420
4.21	QQGGCPGRGTNTCAHSQEELEKYLIRNKCNATVTRFLPLINKVGDNNTTAGNIVSI	480
4.21	QQGGCPGRGTNTCAHSQEELEKYLIRNKCNATVTRFLPLINKVGDNNTTAGNIVSI	480
4.81	GSTEETGKTVIPTSTNGSISNAENSVSQJLSRSSTDSTRALEYTKVKVGANGONAAGPSAD	540
4.81	GSTEETGKTVIPTSTNGSISNAENSVSQJLSRSSTDSTRALEYTKVKVGANGONAAGPSAD	540
5.41	SVTENKIGSPSPKTPVSNVAATSAGSNVYGPBLNSYQOKSSSPFLTRPVYPVPHSENQYFQ	600
5.41	SVTENKIGSPSPKTPVSNVAATSAGSNVYGPBLNSYQOKSSSPFLTRPVYPVPHSENQYFQ	600
6.01	DPRTO1PFEVQYQPOTGYYPPPTVLAGVAPCVPRFRSNVNPESLPPASMPYDHYST	660
6.01	DPRTO1PFEVQYQPOTGYYPPPTVLAGVAPCVPRFRSNVNPESLPPASMPYDHYST	660
6.61	FSPRDRMNSSPYQPPPPQYGPVPPSGMAYAPVYDSSRIWRPMPYQORDD1IRNSLPPM	720
6.61	FSPRDRMNSSPYQPPPPQYGPVPPSGMAYAPVYDSSRIWRPMPYQORDD1IRNSLPPM	720
7.21	DVMHSSVYQTSLRERVNSLDGXXSYACQPPSEPRTTVLPDRPQCHLKTSCBEEQRKRPD	780
7.21	DVMHSSVYQTSLRERVNSLDGXXSYACQPPSEPRTTVLPDRPQCHLKTSCBEEQRKRPD	780
7.81	QWAQHTQAKPLVSLPLVATQSPPLSPLESVDFRADFESSVSGTKFEEDHSHYSPWS	840
7.81	QWAQHTQAKPLVSLPLVATQSPPLSPLESVDFRADFESSVSGTKFEEDHSHYSPWS	840
8.41	CGTIGSCINATDSEPKDVIANSVNLMDLSDGVTKRVHLFETQRRTKEEDP1IPSDGP	900
8.41	CGTIGSCINATDSEPKDVIANSVNLMDLSDGVTKRVHLFETQRRTKEEDP1IPSDGP	900
9.01	IISKNGAISRSRSGTGYHTTDPVQATASQGSATKPIISVSDYVPPYVNAVDSRMSSYGNBEATS	960
9.01	IISKNGAISRSRSGTGYHTTDPVQATASQGSATKPIISVSDYVPPYVNAVDSRMSSYGNBEATS	960
9.61	SAHYVERDRFIVTDLSGHRKHSSTGDLISLQAKNSLILQREANALAMQKWNSLDE	1024
9.61	SAHYVERDRFIVTDLSGHRKHSSTGDLISLQAKNSLILQREANALAMQKWNSLDE	1024
10.21	GRHLTINLLSKETIERENGELQSDYTEDATDKPDRDIELEUSALDDTEPDQSEPEEIL	1084
10.21	GRHLTINLLSKETIERENGELQSDYTEDATDKPDRDIELEUSALDDTEPDQSEPEEIL	1084
10.81	DIQLGISSONDQLINGMAVENGHPYQOQHKEPPKOKQSLGEDHYLEEQTILPPTSCF	1144
10.81	DIQLGISSONDQLINGMAVENGHPYQOQHKEPPKOKQSLGEDHYLEEQTILPPTSCF	1144

QY 721 DWHSSVYQTSLRERYNSLLDGYSVACQPPSEPRTTVLPREPCGHILKTSCSEQI-RKPD 780



FILE REFERENCE: Aeonica-X-1  
 CURRENT FILING DATE: 2001-08-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00671  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00672  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 60/224,687  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 60/2574,203  
 PRIOR FILING DATE: 2001-01-29  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 48290  
 LENGTH: 67  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO ALL136170.2  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
 OTHER INFORMATION: SWISSPROT HIT: Q01970, EVALU 2.60e+00  
 OTHER INFORMATION: EST\_HUMAN HIT: AW962574.1, EVALU 4.00e-32  
 US-09-864-761-48290

Query Match 5.0%; Score 311; DB 9; Length 67;  
 Best Local Similarity 89.6%; Prod. No. 5.4e-14; Indels 0; Gaps 0;  
 Matches 60; Conservative 5; Mismatches 2;

Qy 254 VTKRDEDESSLMLQKEEFFRSYEALRREHDAQIVIWHIAEGLRISPEQWSSLIGDLAHSK 313  
 Db 1 VTKRDEDESSLMLQKEEFFRSYEALRREHDSQIVIAMEAGLRIADQWSSLIGDQSHSKH 60

Qy 314 MOSTIDK 320  
 Db 61 MOSTIDK 67

Query Match 3.3%; Score 207; DB 18; Length 2377;  
 Best Local Similarity 19.3%; Prod. No. 0.00212; Indels 326; Gaps 46;  
 Matches 216; Conservative 152; Mismatches 424;

Qy 48 RACPFDTAINT-----DIDVLPVNPAFLQLVGAQVPDHQSIIKLSNLGEN 92  
 Db 375 KRASFAKSVIQTPEFAPEMYEKIDESVDVYAFGNCMLENATSYV-----421

Qy 93 KHYEVARKCVDELLAKLPLPSGGKGKVASLNOSA-----SRPWNQKLVTLYN 139  
 Db 422 ----YSECQNAAQIYRRVTSGVK-PASFDRKAIPEVKEIEGCIRQNDRYSIKDLN 475

Qy 140 CQLVEEEGRYRAMAARSILGERTTLQHONPQOLSANLWAARARGCOFLGPMQEE 199  
 Db 476 HAFFQEETGVRVELLAEDDGEGKIAKLWLRLIEDDKLK-----GYKDNE 520

Qy 200 ALKLVLIALEDGSALSRKVLVLFVWORLEPRPPQAQSKTSIGHVQ-----LLYRAS 250  
 Db 521 AIEP -----SFDLERDVDPEDAQMVESSYCEDDHTMKAIAKDRVS 563

Qy 251 CPKVTKRDEDESSLMLQKEEFFRSYEALRREHDAQIVIWHIAEGLRISPEQWSSLIGDLAH 310  
 Db 564 LTK-RKREQRQLVREEQEKRKOBESSFKQQNEQQASVS-QAGI----QQLSAASSTQPTA 617

Qy 311 KSHMQSITDKLQSPESFAKSQVQELTIVLQRTGDPANLRPHLEIANDIDPNPDAVSP 370  
 Db 618 PATSAVSTQVPEBEPADQHQ-----LQYQOQPSIVS-----DGTIBSGQGS 662

Qy 371 WEQLENAMVAKTVVQGLVDFIQNTSRKGHT-----POQPQNSKYKTSMCRD 418  
 Db 663 SVFTESRVSQQTVSYG-----SQHEQAHTSTGAQHTVSSIOAQSQPHGVYPPSMSAQ 716

Qy 419 LRQQGGCPRGTRNCTFAHSQEELEKYLURNKKINATVTFPLN-----KVGVNNTVTAGNV 476  
 Db 717 GQNQGQ-PSSSLAGVLSQ-----PIQHPQOOGIQPT-----747

Qy 477 ISVIGSTETTGKIVPSTNGISNAENSYSQLISRSTSTLRALETVKVKGKVANGQNAAG 536  
 Db 748 -----VPSQAV---QYSQAASS-----EG 767

Qy 537 PSADSVTENKIGSPKTPVSNVTAITSAGPSN-VGTELNSVQPKSSPFLTRPVYPHSE 594  
 Db 768 TTAQPVSPQOVSAGTQLPVSQTVATVQGEPHIPVPSVSTQVPHSGAHLPMQOPIPSL 827

Qy 595 NIQYFODRTRQLPFEVQY--PQTGTYPPPPTVPGAVAPCYPRFRSNVPESSLPPASM 652  
 Db 828 LPQY---PVSQ-PISTPHVSTAQTGFSVSPITMAGINQ--PLLTASSATSSIPGSP 882

Qy 653 PYADHST-FSPRDRIKNSPSPY---OPPPPPQPYGPYP-----PVPSPG-MYAPVYD 696  
 Db 883 VVPNQLPTLLOQVNQLOSVQHQLLOPTVOSIG-IPANLQQAEGPLPSGDVLYLQG-FP 940

RESULT 6  
 US-10-840-512-196  
 ; Sequence 196, Application US 10840512  
 ; Publication No. US20050121852A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CAENEPEL, SEAN  
 ; APPLICANT: MANNING, GERARD  
 ; APPLICANT: CHARYDCZAK, GLEN

Qy 697 SRRRIWRPMPYQBDDIIRSNSLSPPMDMHSSVU-QTSURERYNSLGGYSVACQPPSEPR 755  
 Db 941 SR--LPPQYPCDSNTAPSNTASYCHSTVLAAPPMPTEALATOQYFPTVQPYVESTP 997  
 Qy 756 TVPLPREPGHILKTSCEQIRRKPDQWAQYHTQAKLVSS-----TL 797

Db 998 LVPN-GSIVGGQQVQVS-----QPAVSLTQQPTTSQQAVLESSQGVSOAAPPEOT 1046  
 Qy 798 PVATOSPTPPSPPLFS-----VDFRADFSE-----SVSGTKPFEEDHLSHYSPWSSCCGTIG 845  
 Db - 1047 PITOSQTPQPVPLVTSAASAHSDVAGNSDGNENAPSSGSRHEGRTTRKHYRK-- SVRS 1103  
 Qy 846 SCINAIDSEPKDVIAN----SNAVLMLDSDGVKRRYHLFETQRRTKEEDPTLIPPSGP 900  
 Db 1104 RSHBKTSPPKLRLANVKGRVVECOLETNRKMTVKEFDLGDNEBEIATIMVNNDF 1163  
 Qy 901 IISKGKAISRSS-----RTGYHTTDQVQ-----ATASQ 928  
 Db 1164 IL---AERESFVAQRETEIADEMLEDVSVEPESDQGLESLOQ3DDYGFPGSQLE 1219  
 Qy 929 GSATKPIVSVDYVPPVNAVDSRNSSYGREATSSAHYBRDRFRITDLGSHRKISS--TG 985  
 Db 1220 GEFKROPIAVSSMPOOIGVPTS-----SLTOVQHHSAGR-RFIIVSPVPESRRESKVFTS 1271  
 Qy 986 DLIS-LELQAKSNLLIOREANALAMQOKWNLSDEGR 1022  
 Db 1272 DISDPPVASTSQAGMNLSASSLSLQAFSEULKHGQ 1309

---

**RESULT 7**  
 US-10-092-900A-128  
 ; Sequence 128, Application US/10092900A  
 ; Publication No. US20040043382A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Padigaru, Muralidhara  
 ; APPLICANT: Spyttek, Kimberly A.  
 ; APPLICANT: Shenoy, Sureesh G.  
 ; APPLICANT: Toupier Jr., Raymond J.  
 ; APPLICANT: Pena, Carol E.A.  
 ; APPLICANT: Li, Li  
 ; APPLICANT: Zerhusen, Bryan D.  
 ; APPLICANT: Gusev, Vladimir Y.  
 ; APPLICANT: Ji, Weizhen  
 ; APPLICANT: Gorman, Linda  
 ; APPLICANT: Miller, Charles E.  
 ; APPLICANT: Kekuda, Ramesh  
 ; APPLICANT: Patturajan, Meera  
 ; APPLICANT: Gangoli, Esha A.  
 ; APPLICANT: Vernet, Corine A. M.  
 ; APPLICANT: Guo, Xiaojia Sashia  
 ; APPLICANT: Tchernev, Velizar T.  
 ; APPLICANT: Fernandes, Elma R.  
 ; APPLICANT: Casmam, Stacie J.  
 ; APPLICANT: Malyankar, Uriel M.  
 ; APPLICANT: Gerlach, Valerie  
 ; APPLICANT: Liu, Yi  
 ; APPLICANT: Anderson, David W.  
 ; APPLICANT: Spaderna, Steven K.  
 ; APPLICANT: Catterton, Elina  
 ; APPLICANT: Leite, Mario W.  
 ; APPLICANT: Zhong, Haihong  
 ; APPLICANT: Alsobrook, John P.  
 ; APPLICANT: Lepley, Denise M.  
 ; APPLICANT: Rieger, Daniel K.  
 ; APPLICANT: Burgess, Catherine E.  
 ; TITLE OF INVENTION: No. US20040043382A1 Proteins and Nucleic Acids Encoding Same  
 FILE REFERENCE: 214402-290C  
 CURRENT APPLICATION NUMBER: US/10/092,900A  
 PRIOR APPLICATION NUMBER: USN 60/274,322  
 PRIOR FILING DATE: 2002-03-07  
 PRIOR APPLICATION NUMBER: USN 60/274,322  
 PRIOR FILING DATE: 2001-03-08  
 PRIOR APPLICATION NUMBER: USN 60/283,675  
 PRIOR FILING DATE: 2001-04-13  
 PRIOR APPLICATION NUMBER: USN 60/338,092  
 PRIOR FILING DATE: 2001-12-03  
 PRIOR APPLICATION NUMBER: USN 60/274,281  
 PRIOR FILING DATE: 2001-03-08  
 PRIOR APPLICATION NUMBER: USN 60/274,191  
 PRIOR FILING DATE: 2001-03-08

; PRIORITY APPLICATION NUMBER: USSN 60/325,681  
 ; PRIORITY FILING DATE: 2001-09-27  
 ; PRIORITY APPLICATION NUMBER: USSN 60/304,354  
 ; PRIORITY FILING DATE: 2001-07-10  
 ; PRIORITY APPLICATION NUMBER: USSN 60/279,995  
 ; PRIORITY FILING DATE: 2001-03-30  
 ; PRIORITY APPLICATION NUMBER: USSN 60/294,899  
 ; PRIORITY FILING DATE: 2001-05-31  
 ; PRIORITY APPLICATION NUMBER: USSN 60/287,424  
 ; PRIORITY FILING DATE: 2001-04-30  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 768  
 ; SEQ ID NO: 128  
 ; LENGTH: 1267  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-092-900A-128  
 Query Match Score 203; DB 15; Length 1267;  
 Best Local Similarity 19.8%; Prcd. No. 0.0016;  
 Matches 190; Conservative 116; Mismatches 337; Indels 318; Gaps 46;  
 ----- 451  
 Qy 423 GGCPRG--TNCTFAHSSOBELK-----YLRLNKKIN-----  
 Db 21 GRCRNSLYSDC1IEEKTVLQKQDNNEGFGFVLRGAKADTPBEEFTPTPAFPALQYLESV 80  
 Qy 452 -----ATVRIFPLANKVGVNNTVTTAGNVISVI--GSTETTGKIVUPSTNGI----- 496  
 Db 81 DEGGAWAQAGLRTQDFILEVNENNENVKYGHROVNMMIROQGNNHLVLUVTVTRNLDPDT 140  
 Qy 497 -----SNAENNSVSQOLISSTDSTLRALEYTKVKGANGONAGPSADSVTENKI 547  
 Db 141 ARKAAPPKRATTALTRSKSMTSELBDPEE-----VPAKSPS--RAAENNA 191  
 Qy 548 GSPKPTPVSNUAATAGPSN-----VGTEBLNSVPOQKSSPBLTRVPPVPHSEN1QYFQDPR 603  
 Db 192 VEP-----RVATIKQRSRSCRFPGSDMMNVSQRTLGP--RGRCPTVPRLSGJQSVYE-R 243  
 Qy 604 TQIFEVQYPQGQYPPP--TVPAG-----VACPVPRFRSNNVPE 644  
 Db 244 QGIAVMTPTVPGS--PKAFLGIPRGTMQRQKSIGITEEROFQFLAPMMLKFTRSLSPD 300  
 Qy 645 SS-LPPASMPYADHYSTFSPRDRMNNSPPYQOPPPQP--YG-PVPPVPS--GMYAPIYDS 697  
 Db 301 TSBEDIPPPP-----QSVPPSPSPPSPPTYNCPKSPPTPRVYGTQPKAFIQ 344  
 Qy 698 RRIWRPPMQRDDIIRNSNLLPPMDVMHSSVYQTSLRERYNSLDG--YYSVACOPPSEPR 755  
 Db 345 NSAQKVSPTATRSRDTVATM-----MREKGMYFRFELDRY-SLSDSELYSRNAGQFQANFRN 397  
 Qy 756 -TVPLPREFPCGHLKUTSCBQI-----RRKPDQWAQYHTQKAPLVSSLPLPAT----- 801  
 Db 398 KRGQMPENPYSEVGIKIAVYVPAKPARRKGMVWKQSNVEDSPEKTCISIPIPTIVKEP 457  
 Qy 802 -----QSPTRPPSPSPLFSDVPRADFSSESVSGTKEEDHLHYSPWSCT 843  
 Db 458 STSSSGKSGQSSMEIDQAPEPPSQL----RDESLTVS-----SPAAAI 500  
 Qy 844 IGSCINALDSEPKDVIANSNA-VLMLDLSGD-----VKRVRHLPETORRTKEED-- 891  
 Db 501 AG---AVDRERKLEARRNNSPAFLSTQDGEDVGLGPAPRTRPSMPBEGDADEDA 556  
 Qy 892 -----P11PFSDGII1ISKNGAISR-----SSRTGYHTTDPVQATASQGSA-- 931  
 Db 557 EQLSPMPSATPRBEPHFVGAEEAARGEGRPLNSTSKAQGPESSPAVPSASSGTAGP 616  
 Qy 932 -----TKPISV-----SPDVP-----YVN-----AVD 948  
 Db 617 GNYVHPLTGRLLDPSPLAALSARDRAMKESOGPKGEAPKADLNKPBYIDTMKRPSDL 676  
 Qy 949 SRNSSYGNBET-SSAHYVERDRFTVTDLSGHRKHSSTGDL--SLELOQAKNSNLLQRE 1005  
 Db 677 AGFTTVTRONTGRQRLRQETENCYETDUGDRKQDDKQMLIDIMDTSQKSACLLMVHT 736

Query 1006 ANALAMQQKWNSLDEGRHLTLNLSSKEELRNGELOQSDYTEDATDTK-  
Db 737 VDTKLKDNLQEBE-----KAEVEMKDSSPSSVPEGVSETEGALQISAAAPPPT 787  
Query 1053 -PDRDI----ELE-----L-SALDDDEPDQSEQEELDILQLGISSONDQLL 1094  
Db 788 VPERTIVAVGSMEAVILFRIPPPPLASAVLDDEFIFTEPLPPPLEFANSFDIPDRAA 847  
Query 1055 NGMAVENGHPVQOHQKEEPKQRKSLSGEDHVILLEBQTKTILPVTSQPLPVSIASACL 1154  
Db 848 SVPAL-----SDLVKQKSDTFQSPSLNNSQ---PTNSADSKK-PASLSN-CL 890  
Query 1155 P 1155  
Db 891 P 891

RESULT 8  
Sequence 6679, Application US/10732923  
Publication No. US20050108791A1

APPLICANT: Edgerton, Michael D  
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES  
FILE REFERENCE: 38-15 (52796) C  
CURRENT APPLICATION NUMBER: US/10/732,923  
CURRENT FILING DATE: 2003-12-10  
PRIOR APPLICATION NUMBER: 10/310,154  
PRIOR FILING DATE: 2002-12-04  
NUMBER OF SEQ ID NOS: 24149  
SEQ ID NO 6679  
LENGTH: 728  
TYPE: PRT  
ORGANISM: Candida tropicalis  
US-10-732-923-6679

Query Match 3.2% Score 200.5; DB 17; Length 728;  
Best Local Similarity 20.1%; Prod. No. 0.0001; Mismatches 29; Indels 225; Gaps 36;  
Matches 158; Conservative 108; Mismatches 29; Indels 225; Gaps 36;

Query 396 SRKGHETPOPOPNSKYKTSMCRDLRQGCCPRGTNCFAHSQEELEYKVRLRN-----447  
Db 2 SKCSNSTNRP-----GDPRVTKNAFVH-----KLYTMNDPNLSHLI 38

Query 448 -KKINATVTRTPFLNKVGVNNTTT-TAGNIVSIGSTETGKIVPSTNGISNAENSV 503  
Db 39 WWTNNNVBNNTFLALYPGKEFANCILTRYFKHGNVASFYRQLHNYG---FHKVSDPHNNS 93

Query 504 SOLISRSTDSTURALETVKVGANGONAAGPSADSTENKIGSPP-----551  
Db 94 SS-----SALSINHNSYSMNTVHQHNNAGG-----AHNNKDVPPIWEFKHSSGKFK 143

Query 552 --KTPVSNVAATSAQPSSNGTEINS-----VPOKSSSPFLTRVY-----Y 589  
Db 144 KGDEKSLALIKRSSSSNSGSRNNTNETLTLQNHVHPVPPQO--PMXQEYPIPQONITY 201

Query 590 PPSENIIQYFQD---PROQIPF---EVPOYPOTGYPFPPTVPGVAPC-VPRPVSN- 640  
Db 202 DPYNLNNFQQQPPPPQPMYNNCNPQ-PVPPPLFPPLAANQQLPPQFYGGQP 260

Query 641 --NVPESLPPASMPYADHYSTSPDRMNNSPYQPPPPQGPVPPPSGM-----690  
Db 261 YFQYPGSLIPPQSQP-----QPVPPQPPQ--AAAPPSPGTYQFINTFSQ 302

Query 691 -----YAPVYDSRRIW-----RPPMFTD---DITRSNSLPPMD 721

Query 703 FPPSPSPQQPFIQMNLLQSQNSASNTPQIQQPTPQHVEBPTRKQDCITDTHVKELESRSRSP 362

Query 775 IRRKPDDOWAQVHTQAKPLVSSTLPVATQSP----TTP-----SPLFSYDRAFTSES 822  
Db 423 ILPPPSOL-----NRSSSLNSYPLPTSSQPSIMAAPPPEMRRTSTPLU-----HS 469

Query 823 VSGTKEEE----DHLSHYSPWMSCGTIGSCINAIDSEPKDVTIANSLAVMLMDSG---DV 874  
Db 470 LSHSNTNGNTTSRLSNSIPARNSPLSSTAESTTPNIVSSSSQPSSTSSGGLNIPNV 529

Query 875 KBRVHLEFQTORPTKEERDPIIPPSDGPIISKNGAISSRSRTGYHTDPVQATASGATKP 934  
Db 530 TRKPSLFSSSLQERLRPSVFDTIH-----NENHOSGIKSPTIQTYSKNS--- 576

Query 935 ISVSDVPPVNAVDWSRMSSYNEATSSAHTYVERDRPIVTLI-SGRKHKSTGDLLSLEQ 993  
Db 577 ----IVSNTSISIFSKSSISIASTPNQHGANNRPSLSSITNSVKEEEVKEEIV 631

Query 994 QAKNSNLLOREAN-----ALAMQOKWNSLDEGRHLTLNL---SKIELRNGELQ 1041  
Db 632 IKNGKSLANDNSKSWTPMNGNASVQDREQRSLTESPNLKSEICVOPENKRVSV-NSLX 690

Query 1042 SDYTEDAT 1049  
Db 691 DDDNEDKT 698

RESULT 9  
US-10-334-143-77  
Sequence 77, Application US/10334143  
Publication No. US20040009549A1  
GENERAL INFORMATION:  
APPLICANT: GRIGORIEV, IGOR VYACHESLAVOVICH  
APPLICANT: SUDARENANAM, SUCHA  
TITLE OF INVENTION: METHOD FOR DETECTING REMOTE HOMOLOGUES AND NOVEL  
TITLE OF INVENTION: KINASES IDENTIFIED WITH THE METHOD  
FILE REFERENCE: 038602/1543  
CURRENT APPLICATION NUMBER: US/10/334,143  
CURRENT FILING DATE: 2001-12-31  
PRIOR FILING DATE: 2002-12-31  
PRIOR APPLICATION NUMBER: 60/343,169  
NUMBER OF SEQ ID NOS: 207  
SEQ ID NO 77  
LENGTH: 3586  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-334-143-77

Query Match 3.2% Score 198.5; DB 15; Length 3586;  
Best Local Similarity 19.7%; Prod. No. 0.0017; Mismatches 478; Indels 409; Gaps 64;  
Matches 263; Conservative 187; Mismatches 478; Indels 409; Gaps 64;

Query 72 QLVGQVPDFQHSIKLNLGENKHYEVAK-----KCVDIAL-YLKPLSGKG 117  
Db 1421 EVIGGDYKRTIHNLLSTSNNKISERAKIDASERGNVQFFTCIEAGDALYKQLHL-----1476  
Query 118 VASLNQSALSRMQRKLTVLNCQLVVEEGVRMARAARSGLERTVTELIL---OHQ-- 171  
Db 1477 -TESNETLTAKQGEKEIIGG---DVEGKMLKROSLYVERTSETD1PGDVHNTV 1531  
Query 172 ----NPQQLSANLWAAVRARGCQFLGPAMOEAKLVLALED--GSALSRKVL----- 219  
Db 1532 KVMTEPQST-----FGKPKSEIKIKGDLTSLNSLSQAVNQKTVKTE 1576

Query 220 ----VLFVVQBL-----EPRFPOASKTSIGHVYQOLLYRASCK-----VTKRDE 259

Db 1577 IKGNMNLATLKSILKESSHWRKESKQPAIPGDETERCLEKATNTKTEILKKELLKDDL 1636

Query 260 DSMLQKKEEFPSYEARRE---HDAQIVHIAAMEAGLRISEQWSSLYGLDAHKSHMQ 315

Db 1637 ETSLSLKEAQAQSFFKEYKEGVTKDOK---AVMAG--SSGEQKTDI---HQAVQERNKN 1688

Query 316 SIIKDQSP-ESFAKSVOELTIVLQRTGDPANLNRPHL-----LLANIDP 362

b	b	1689	SLLQPKPGPPEPAAKWQGADTLSCSTMGSCHGNLVEERTEVNLPKAPKGTKIVIIREQ	1748
Y	b	363	NPDAYSPTWEOLENA-----	-MYAVKTVVHGLYDFIONY-SRKGHETPOPOPNSK 410
Y	b	1749	NNDALEKSLLRLSNSHKSNVLESSDXTKGVWTDTGEQLRDEMSROLTSTSVSNNLT	1808
Y	b	411	YKTS-MCRDLR-----	-QGGCPRGPNCTF-AHSOEELEYKRBNKKINATVRTP 458
Y	b	1809	TKESEDRARELKDDOVENQSAGRTVGKQTYELRNHDQKMEGFHISPKCKTNK---	1865
Y	b	459	LLNKYGVNNTVTITAGNVISGTEBTGKIVPNTINGTSNAENSVSQSLRSSTDSTURAL	518
Y	b	1866	-----	-----
Y	b	519	ETV---KKGVKGVGANGCON-----	-ARGPSADSVTENKIG-----
Y	b	1905	ETKYSNKKDIKCKNNINLQPMWQQLPVEDTSNVTENKVSEKSHNTFKATNKKRETDVHLKS	1964
Y	b	549	-----	-SPPKTCPVSNVTAATSAGPSNVGTTELNSVPOKSSPLTRVPVYVPHSENIQYQDP 602
Y	b	1965	QDFLMKMTNTSTGLKWAEMERSLNPNFNPB-----	-NNVYSECECPPLPPSPPPPS-----NA 2016
Y	b	603	RTQIPPEVPOQYQTYGYPPPTPVAPGVACPVPRFRS-----	-NNVPESSSLPPASM 652
Y	b	2017	SSEIEBPLP-----	-PPPLM---MFPEKNGFTPLSSTEKIKAEEFSFGJPLPP--- 2062
Y	b	653	PYADHYSTFSFRDRMNNSPQYQQPQYGPVPPVSYGMAYVYDSSRRIWRRPMYQRDDII	712
Y	b	2063	PPVDEKS-----	-ERSSSMFLPPPP-----PPTPS-----QKPAHL 2093
Y	b	713	RSNSLSPMDVMHSS---VYCQPSLREYNSLGDGYYSV-----	-ACOPPSBERTVPLPREPC 764
Y	b	2094	LSSSAPE--KHSGIDMQYTSQKEASLNSNOSQAKitGKGTVGLVLP-----	-TLPKPKLP- 2144
Y	b	765	GHLK-----	-TSCEBQIRRKPDQ-----WAQHTQ-KAPLYUSTLKVAT 801
Y	b	2145	KHIKONKNDSPKVKVLAETSLSDMECKITSKDKQKVMWTSSEBTETKVNVLKSILERK	2204
Y	b	802	QSPTPPSPLFSEVDFRSSEVSGXKFEBH-----LHSYSPMSCGTGSCINAIDSBP	855
Y	b	2205	Q-----LSIDSAACLSHTVPGT SAPRKQKIAPILKSHSPFSSG-----QONP 2247	
Y	b	856	KDVIANSNAVMLDLDGDKVRVHLFETOQRTKEPDPLPFSODPPIISWKGA-SRSRSSTG	915
Y	b	2248	KPYMRKFKTPML-----IAEEKYRQOQEIEEIKQKQ-----	-SSYYNIVKTSQSN 2291
Y	b	916	YHTTD-----PVQATASQGSATKPISVSDVYPTYNIAVDSRWSSYGN-----	-AT 959
Y	b	2292	QHITEVEKEMPLQKT-----NEETSLSGIDSECTVQBPSPGSNSNARLIGVCSDNQNLST	2345
Y	b	960	SSAHYTERDRFIVTULSGHRRKHSSTS DLL-----SLEJOQAKNSNLJO--REANALAMOOK-	1014
Y	b	2346	TSPETVAAKRLHHVLAESDKRMKKEVLOSSRDMQSKSACEIKOSHOBECSTQTQQJKK	2405
Y	b	1015	--WNSLDEGRHILTMILKSEIELNGELOSDDTAD-----TKPDRDIEIL-----S	1062
Y	b	2406	YLBQLHPQSPKSPAFKVKTLPT----LDHNTNETDHSYBKHQSBIDQPTFTKQ	2461
Y	b	1063	ALDTPEDDGQSEPBEILDIOLGISSQNQDQILNGMAVENGHPVQHQKEPPPKOKK-----	1117
Y	b	2462	YLTKKCFEASTE-----	-CSHKQSK-----ABHYOLPKKEKRVTVQL 2498
Y	b	1118	--QSLGEDHVILEBQRTILPVTCSPQSLPVSIASHASCLPITSVSAGNLIL--KTHM 1172	
Y	b	2499	PTEISQRNQ--EDKLKMWV-----RKOREFSGSORGKLPGSEEKNQGSMIGKEERIJ	2550
Y	b	1173	SEDKNDLTKPVANGOMV 1189	
Y	b	2551	TERKEHILKNGSAPVV 2567	

RESULT 10  
US-10-437-963-126940

Qy 976 SGHRKHSSTG---DLSLEOOAKNSLLOREANALAMOOQKWNSLDEGRHLTNULSK 1031  
 Db 868 - ARQRPVAAKVKLLIDQLQRSKNEVIMNIKMPLP-----DIMNSVIAL 913

Qy 1032 EIELRNGELQSDYTEDATDTKPDRI 1057  
 Db 914 DDSIVDG-QvdylKECPTKSEMEI 938

---

RESULT 11  
 US-0-864-761-42017  
 Sequence 42017, Application US/09864761  
 Patent No. US20020048763A1

GENERAL INFORMATION:  
 APPLICANT: Penn, Sharron G.  
 APPLICANT: Rank, David R.  
 APPLICANT: Hanzel, David K.  
 APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR FILE EXPRESSION ANALYSIS BY MICROARRAY

CURRENT APPLICATION NUMBER: US/09/864,761  
 CURRENT FILING DATE: 2001-05-23  
 PRIOR APPLICATION NUMBER: US 60/180,312  
 PRIOR FILING DATE: 2000-02-04  
 PRIOR APPLICATION NUMBER: US 60/207,456  
 PRIOR FILING DATE: 2000-05-26  
 PRIOR APPLICATION NUMBER: US 09/632,366  
 PRIOR FILING DATE: 2000-08-03  
 PRIOR APPLICATION NUMBER: GB 24263.6  
 PRIOR FILING DATE: 2000-10-04  
 PRIOR APPLICATION NUMBER: US 60/236,359  
 PRIOR FILING DATE: 2000-09-27  
 PRIOR APPLICATION NUMBER: PCT/US01/00666  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00667  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00664  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00669  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00665  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00668  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00663  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00662  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00651  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: PCT/US01/00670  
 PRIOR FILING DATE: 2001-01-30  
 PRIOR APPLICATION NUMBER: US 60/234,687  
 PRIOR FILING DATE: 2000-09-21  
 PRIOR APPLICATION NUMBER: US 09/608,408  
 PRIOR FILING DATE: 2000-06-30  
 PRIOR APPLICATION NUMBER: US 09/774,203  
 NUMBER OF SEQ ID NOS: 49117  
 SOFTWARE: Annotrax Sequence Listing Engine vers. 1.1  
 SEQ ID NO 42017  
 LENGTH: 123  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 OTHER INFORMATION: MAP TO AL136170.2  
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.9  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.4  
; OTHER INFORMATION: EST HUMAN HIT, NS9624.1, EVALUATE 2.00e-24  
; OTHER INFORMATION: SWISSPROT HIT: Q27409, EVALUATE 1.10e+00  
 US-09-864-761-42017

Query Match Score 195.5; DB 9; Length 123;  
 Best Local Similarity 35.8%; Pred. No. 1.4e-05;  
 Matches 48; Conservative 20; Mismatches 43; Indels 23; Gaps 7;

Qy 618 YPPPPPTVPGAVAPCPYPRFRSNNNPPESSLPPASMYADHYSPDRDNSSPPYTOPPPP 677  
 Db 2 YTPPP-----QCYSRFV-----PPPSADEPAPPYLHDHYPPYL\_QERVNSQYCTQPO 49

Qy 678 QPYGPYBPPVPSGMAYAPVDSRRIRPPMYYQDDITRSNLLPPMDYHSSV--YCTSLRER 735  
 Db 50 Q---YFPI---YPSHYDGRRVYAPSYTREEIFPRESPI-PIEPPAAVSYPPESRER 100

Qy 736 YNSLDGYSYACQP 749  
 Db 101 YQQIESSTYPVAPHF 114

RESULT 12  
 US-10-288-798-9  
 Sequence 9, Application US/10288798  
 Publication No. US2003020729A1

GENERAL INFORMATION:  
 APPLICANT: BANDMAN, Olga; NGUYEN, DannieL B.;  
 APPLICANT: WALIA, Narinder K.; HAFLAJA, April J.A.;  
 APPLICANT: YAO, Monique G.; GANDHI, Ameena R.;  
 APPLICANT: GURURAJAN, Rajegopal; DING, Li;  
 APPLICANT: PATTERSON, Chandra; YUE, Henry;  
 APPLICANT: BAUGHN, Mariah R.; TRIBOLEY, Catherine M.;  
 APPLICANT: THORNTON, Michael; ELLIOTT, Vicki S.;  
 APPLICANT: LU, Yan; ISON, Craig H.;  
 APPLICANT: AU-YOUNG, Janice; TANG, Y. Tom;  
 APPLICANT: MARCUS, Gregory A.; ZINGER, John D.;  
 APPLICANT: LU, Dyung Aina M.; LAL, Preeti G.;  
 APPLICANT: RAMKUMAR, Jayalaxmi; WARREN, Bridget A.;  
 APPLICANT: KEARNEY, Liam; POLICRY, Jennifer L.;  
 APPLICANT: THANGAVELU, Kavitha; BURFORD, Neil  
 TITLE OF INVENTION: HUMAN KINASES  
 FILE REFERENCE: PI-0209 USA  
 CURRENT APPLICATION NUMBER: US/10/288,798  
 PRIOR FILING DATE: 2002-11-01  
 PRIOR APPLICATION NUMBER: PCT/US01/27219  
 PRIOR FILING DATE: 2001-08-31  
 PRIOR APPLICATION NUMBER: US 60/240,542  
 PRIOR FILING DATE: 2000-1-13  
 PRIOR APPLICATION NUMBER: US 60/238,389  
 PRIOR FILING DATE: 2000-10-06  
 PRIOR APPLICATION NUMBER: US 60/236,499  
 PRIOR FILING DATE: 2000-09-29  
 PRIOR APPLICATION NUMBER: US 60/234,902  
 PRIOR FILING DATE: 2000-09-22  
 PRIOR APPLICATION NUMBER: US 60/232,654  
 PRIOR FILING DATE: 2001-01-14  
 PRIOR APPLICATION NUMBER: US 60/231,357  
 PRIOR FILING DATE: 2000-09-08  
 PRIOR APPLICATION NUMBER: US 60/229,873  
 PRIOR FILING DATE: 2000-08-31  
 NUMBER OF SEQ ID NOS: 48  
 SEQ ID NO 9  
 LENGTH: 2135  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US2003020729A1 5502218CD1  
 US-10-288-798-9

Query Match 3.1%; Score 195; DB 15; Length 2135;  
Best Local Similarity 19.1%; Prod. No. 0.0013; Mismatches 528; Indels 378; Gaps 56;  
Matches 254; Conservative 169;

Qy 114 GIGKVASINQNSALSRPMQRKLTVNLNCOLVEBGRVRAARSGLGERTVTTELILOHONP 173  
Db 720 QDPSSSTLTGVSQQPIRPOQQGIOQTAPPOTVQYLSQSQTSSSEATAQPVSQQAP 779  
Qy 174 QQLSANLWVAARVRGCOLGPA----MQEBAKLVLIALEGDSALSRKVLYFVYQRL 228  
Db 780 QVPQPQVSAGKQSTGVSQVAPSPVAVQAPQATQPTTAASSH----VASGNSDGN 836  
Qy 229 PRFPQASHTSIGHVQLLYRASCFKVTFRDEDSLMLQK----EFSRSTBALRRHDA 282  
Db 837 ENVPSSSRHEGRTRKRYRKVS-VRSRSHREHTSRPKRILNVNSNKGDRVVCEOLETHNR 895  
Qy 243 QYTHIAMBAGRLSPEQWSSLY-GDLAHKSHMQSIIDKLOSPESPARSVQELTIVLQRT 341  
Db 886 KMYTFKFILD-GDNPEELATIMVNDFLAIREFSTFDVQRETIKEADEMSEDVSYPE 954  
Qy 342 GDPANLNMRPHIELAND----PNPDAVSPPTWOLENAMVAVKVV 385  
Db 955 GDQG----LESLQKDDYFGSGSCKLGEFKQPPASMP----OCIGIPTSSUTQVV 1004  
Qy 386 HGLYDFINQYSRKGHETPOPQN----SKYKTSMCRDLRQGGCPRGNTCTPASQ 437  
Db 1005 H----SAGRFTIVSPVPESTRRESVFPSEITDTVASTAOSPG----MNLSHSA 1051  
Qy 4348 EELEYKR----LRNNKKI----NATVRTRFP----LINKGVVNNTTTA----473  
Db 1052 SSSLQQAFSELRAQMTEGPNTAPPNHSHTGTPVPPFLISIAGPTTAATAAPVA 1111  
Qy 474 ----GNTVISVGTSSTETTKIVPSTNGNSAENS----VSQIISR 509  
Db 1112 TSSPPNDTSV1QSEVT----VPTEGIAGVATSTGVTSQGGHPIPVSSEPVLSYSS 1168  
Qy 510 STPSTLRALETVKVKGKGAN----CQNAAGPSAD-SVTENKIG----SPBKTPVSN 558  
Db 1169 ITIPAVVSISSPSLQPTSTSEIVVSTALYSPVTSATSASAGSSTATGPKPAV 1228  
Qy 559 AATSGAGPSVNGTBLNSV----POKSSPFTRVPPHSENTOYFQ 600  
Db 1229 SQOAGSTVTGATLTSVTTSPSTASQLS1QSSSTSPTIAETVVAHSLD-KTSH 1287  
Qy 601 DPFTQIPEPVPOVPTGYPPPP----TVPGAYPCY-PFVRSNNVPESSUPASMPY 654  
Db 1288 SSTSTGLASH-SAPSSSSPGAGVSSYISQPGSLIHPVLPVTSVIASTPLQQAQPTSTPL 1346  
Qy 655 ADHYSTFSRDRMANSSPQOPPPQPYGVPPVPSGM----YAPVYDSR 698  
Db 1347 ----LPQVPSIPLVQPVANVPAQVTLIHSQPAALLNPQTHCPEDSD 1394  
Qy 699 RIWRPPMQRDDI----IRS--NSLPPMDYMHSSV-YQTSLSRERYNSLDGXSYA---746  
Db 1395 T--QPKAPGIDDKTLEKSLRSLFSEHSSGAHASVSYLTSLVTESTVTPGIPTAVAP 1452  
Qy 747 ----CQPSE-PRTTVPLPREBC--GHJLKTSEEQIRR-KPDQWQYHFTOKAP 791  
Db 1453 SKLLTSTSTCLUPTNLPLGTVALPVTPVTVPEQVSTVSTTSGVRKGT----APSKPP 1508  
Qy 792 LVSS-TLVAVATQSPPTPPSPLFSVDFADESESVSGTKFEDHLHYSPWSCTGSCINA 850  
Db 1509 LTKAPVPLVPGTTEL----AGTLPS-EQ 1530  
Qy 851 IDSEPKDVTSANAVLMQLDS----GDVKRKRVLHFETORRTK----888  
Db 1531 LPPPPGPSPLTQSOOPLDIAQLRRTLSPEMITVTSAVGPVMSMAAPTAITEAGTQPKGV 1590  
Qy 889 --BEDP11PFSQDPIIIKSKG----AISRSRTGYHTTDPVQATASQGSATKPIVSVD 939  
Db 1591 SQYKEGPVLTSSAGVPMGRQFQVSVAADGAQEGKNSKSEDASVHFESESSVLISS 1650

Qy 940 YVPYNAV-----DSRWSSYGNATE-SAHYVERDRFIVTDLS---GH 978  
Db 1651 SSPESTLVKPEPNITIPGISSDVPPEAKTITASEAKTSDFGOPTKVERFQVTTANKVGR 1710  
Qy 979 RKHSSTEDLIS----LELOQAKNSNLLQREANALAMOQKN----1016  
Db 1711 FSVSKTEDKIDTKKEGPVASPPEMDLEQAVLPVAPIKKEPELSEPHLNGGPSDPEAA 1770  
Qy 1017 ---SLDBGR---HLTUNLSSKEIERN----GEIQSDY-TEDATDTKPDRIELELS 1062  
Db 1771 FLSRDVDDGSGSPHSQHOLSSKSLPSQNLSQLSNSNSNSMISSDNESDIEBEDLKLER 1830  
Qy 1063 ALDDEDGQSSPTEBIDIOLGISSONDOLLN----GMAVENHGPVQHQKEPPKQ 1115  
Db 1831 RU----RDGHKLKEQDLQERQRKHIESLYTKLGKVPVAVIUPPAPLSSRRRPTKS 1883  
Qy 1116 K----KQSLGEDH-----VILEEQKTILPVTSCSFQQLPVTSNSACSLPI 1156  
Db 1884 KGSKSSRSSSLGNKSPQLSGNLGGQSAASVLUHPQQTLLHPQGN----IPESCONQJQPL 1938

Qy 1157 TTSVAGNL 1165  
Db 1939 KFSPSSDNL 1947

RESULT 13  
US-10-362-892-9  
Sequence 9, Application US/10362892  
Publication No. US200403888A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.; BANDMAN, Olgia  
APPLICANT: NGUYEN, Daanil B.; WALIA, Narinder K.  
APPLICANT: GANDHI, Amena R.; GURURAJAN, Rajagopal  
APPLICANT: DING, Li; PATTERSON, Chandra S.  
APPLICANT: YUB, Henry; BAUGHN, Mariah R.  
APPLICANT: TRIBBLEY, Catherine M.; THORNTON, Michael B.  
APPLICANT: ELLIOTT, Vicki S.; LU, Yan  
APPLICANT: ISON, Craig H.; AU-YOUNG, Janice K.  
APPLICANT: TANG, Y. Tom; AZIMZAI, Yalda  
APPLICANT: BURILLI, John D.; MARCUS, Gregory A.  
APPLICANT: ZINGLER, Kurt A.; LU, Duyng Aina M.  
APPLICANT: LAL, Preeti G.; RAMKUMAR, Jayalaxmi  
APPLICANT: WARREN, Bridget A.; KEARNEY, Liam  
APPLICANT: POLICKY, Jennifer L.; THANGAVELU, Kavitha  
APPLICANT: BURFORD, Neil  
TITLE OF INVENTION: HUMAN KINASES  
FILE REFERENCE: PF-0209 USN  
CURRENT APPLICATION NUMBER: US/10/362,892  
PRIOR APPLICATION NUMBER: PCT/US01/27219  
PRIOR FILING DATE: 2001-08-31  
PRIOR APPLICATION NUMBER: US 60/229,873  
PRIOR FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: US 60/231,357  
PRIOR FILING DATE: 2000-09-08  
PRIOR APPLICATION NUMBER: US 60/232,654  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US 60/234,952  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: US 60/236,499  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: US 60/238,389  
PRIOR FILING DATE: 2000-10-06  
PRIOR APPLICATION NUMBER: US 60/240,542  
PRIOR FILING DATE: 2000-10-13  
NUMBER OF SEQ ID NOS: 48  
SOFTWARE: PERL Program  
SEQ ID NO 9  
LENGTH: 2135  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: misc\_feature  
 OTHER INFORMATION: Incyte ID No. US20040038881A1 5502218CD1  
 US-10-362-892-9

Query Match Score 195; DB 15; Length 2135;  
 Best Local Similarity 19.1%; Pred. No. 0.0013; Mismatches 528; Indels 378; Gaps 56;

Matches 254; Conservative 169; MisMatches 528; Del 378; Gap 56;

114 GKGVAISNQASLRPMQRKLVTLVNCOLVEEGRVAMRAARSLGERPTVTELLHQNP 173  
 720 GQPSSSSTLGVSQQPQHPOQGQIOTAPPQTVOYLSQTSOTSTSPPATAQPVSQFQAP 779

174 QQLSANLMAAVRARGCQLGPA----NQEAEALKVLIALEDGSALSRKVLFVIVQRL 228  
 780 QNLQVSGASKQSGVQSPVVAAPPVAYAQPTOPTIASSUDSAHSD--VASCHSDGN 836

229 PRFQASTKSIGHVQLVYLARSCFKVTRDESSLMLRK----EEFRSYEARREHDA 282  
 837 ENVFSSRGHEGTTKRYRKS-VRSRSHKEKTSRPKRILAVSNKGDRVVCEOLETHNR 895

283 QYHIAFAGLRSPEQWSSLY-GDLAHKSHMQS1DQKQSPESFAKSQVQELTIVLQT 341  
 896 KMYTFKFID-GDNPEETATIMVNNDFLIAERESFVTDQVRETIKEADEMISSEDVYPE 954

342 GDPANLNRLRPHIELLAND----PNPDAYSPTWEQLENNAMAVAKTVV 385  
 955 GDQG----LESLQKDDYFGSGSKQLEGFKQPAASSMP--QCIGIPTSSU7QVV 1004

386 HGLVDFIQNSYRKGHETPOPQPN----SKYKTMCRDLRQGGCPRGNTCTFAHSQ 437  
 1005 H----SAGRFIVSPVPSRLEVISVFPSEITDTAASTAQSPG----MNLSHSA 1051

438 EELFKYR---LANKKI----NATYRTEP----LINKYCUUNNTYTAA----473  
 1052 SSISSLQOAFSELRAQMTEGPNTPAPPNHSHTCGTFPVVPPFLSSIAGVPTTAATAPVA 1111

474 ---GNTISVIGSISTETGKIVPSTNGSNAENS----VSOLISR 509  
 1112 TSSPPNDTSVIOSEVT--VPTEEGAGVATGCVYVTSGGJPIPVYSESPVLISSVY 1168

510 STDSTLRALETVKVKGAN----SONAAGPSAD-SVTENKIG---SPBKTPTVSNV 558  
 1169 ITIPAVVSISSTPSLQPTSTSEIVVSTALYPSVTATSAGSSTATGPKPAVV 1228

559 AATSGAPSNGTFLNSV----POKSSPELTJVPUVYPHSENIOYFQ 600  
 1229 SQQAGSTGTVGATLTSVSTTSPSTAQSIOSSSSSTPTIAETVVSANSLD-KISH 1287

601 DPRTOQIPFEPVQYQPTGYPPPP----TVPGAYAPCY-PFRYRNNVNPESLPPASMPY 654  
 1238 SSTTGFLSPU-SAPSSSSSPGAGVSSYISOPGSHIHPVIPSVIASTPLPQAGPSTPL 1346

655 ADHYSTFSPRDRMNSSPYQOPPPQPYGVPPGPM-----YAPYVDSR 698  
 1347 -----LPQVPSITPLPQVANVPAVQOTLHSQOPALLPNQPHTHCPEDSD 1394

699 RIWRPPM10RDI----IRS--NSLPPMDVNHSSV-YOTSLRERYNSLDGYYSVA---746  
 1395 T--QPKAPGIDDKTLERLSLSEHSSGAGHSVSYLSETSIVESTVTPGJPTTAVAP 1452

747 -----COPPSE-PRTTVPPLPREC---GHLKTSCEEQIRR-KDQMAQYHTOKAP 791  
 1453 SKLJSTSTSTCLPPTNLPIGTVVALPVYVTPGQVSTPVSTTISGVFGT---APKPP 1508

792 LVSS-TLPVATOSPTPPSPLFSVDFADESESGTKFBEDHLSHYSPPCSGTTGSCINA 850  
 1509 LTKRKPVLPGTEL-----AGTLPS-EQ 1530

851 IDSEPKDVTANSAVLMQDLS----GDVKKRVRHLFETORTK---888  
 1531 LPPPGPSITQSOQPLEDDAQLRRTLSPEMITVTSAGPVMSMAAPTITEGTOPKGV 1590  
 889 --EEDP1IPPSDGPIISKWG----AISRSSRTGYHTDPVQATASQGSATKPIVSD 939

1591 SQKEGPVLATSSGAGVFKMRFQVSVAAADGAKQEKRNKSDAKSVHFESESSTSESSVLS 1650  
 940 YPVYNAV-----DSRWSSYGNBEATS-SAHYVERDRFIYTDLS---GH 978  
 1651 SSPESTLVKPEEPNGITIPGISSDVPESAHKTTASEAKSDTCOPTYKGRFOVTTANKVGR 1710  
 979 RGHSTGDLLS----LELOQAOKNSNLQREANALAMQOKWN-----1016  
 1711 FLSRVDNGSGSPHSQHQLSSRKLSPONLSOSLNFSNSYMSNSDIEDDLKLER 1830  
 1017 ----SLDEGR---HLTINLNSKIEBLRN----GELOSDY-TBDATDTKPDRIEELS 1062  
 1771 FLSRVDNGSGSPHSQHQLSSRKLSPONLSOSLNFSNSYMSNSDIEDDLKLER 1830  
 1063 ALDTEBPDGQSBPIEELDOLGISSONDOLLN----GMAVENGHPVOHQKEPPKQ 1115  
 1831 RL----RDGHLKEQDLQSRQHIELESLYTKLGKVPAPVIIPPAAPSGRRRRETKS 1883  
 1116 K----KQSLGBDH----VILEBKTHLPVTSFCSFQPLPVTSNSACSLPI 1156  
 1884 KCSKSSRSSSLQNKSPOLSGNLSGOSAASVLUHPQQTILHPPGN----IPESQGNQOLLQPL 1938  
 1157 TTSVAGNL 1165  
 1939 KPSPPSDNL 1947

RESULT 14  
 US-10-196-935A-2  
 Sequence 2, Application US/10196935A  
 Publication No. US0030082720A1  
 GENERAL INFORMATION:  
 APPLICANT: Lifton, Richard P  
 APPLICANT: Wilson, Frederick H  
 APPLICANT: Choate, Keith  
 APPLICANT: Ishikawa, Kunihiko  
 APPLICANT: Nelson-Williams, Carole  
 TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO TREATING AND DIAGNOSINC  
 TITLE OF INVENTION: HYPERTENSION  
 FILE REFERENCE: 044574-5113  
 CURRENT APPLICATION NUMBER: US/10/196, 935A  
 CURRENT FILING DATE: 2002-10-25  
 PRIOR APPLICATION NUMBER: US 60/306, 084  
 PRIOR FILING DATE: 2001-07-17  
 NUMBER OF SEQ ID NOS: 6  
 SEQ ID NO 2  
 LENGTH: 2382  
 TYPE: PRT  
 ORGANISM: Homo sapiens

US-10-196-935A-2

Query Match Score 195; DB 14; Length 2382;  
 Best Local Similarity 18.5%; Pred. No. 0.0015; Mismatches 506; Indels 362; Gaps 49;  
 Matches 235; Conservative 167; Mismatches 506; Indels 362; Gaps 49;

48 KACPFDQTAINT-----DIDVLPVNFAFLQLVGAQVDHOSIKLNSLGEN 92  
 375 KRASFAKSVIGTPPEWAPMEYBEKYDESDVYAFGMCMLEMATESEY-----421

93 KHYEVAKKCVEDIALYLKPLSGKGVASLNQSL-----SRPMORKLVTLVN 139  
 422 ----YSECQNAQAIYRRTVSGVK-PASFDKVAIPEYKEIIEGCIQRNQKDERYSIKDLN 475  
 140 CQLVEEEGRVRAARSGLGERTVTELJLQHONPQQLSANLWAARARGCQLGPAMQBE 199  
 476 HAPFQETGVRYVELABEDGEKIAKULWLRIDIKLK-----GKYKDNE 520

200 ALKVLLAEDSALSRSKVLVLFVQLEPRPQASRSTSIGHVULLYRASSFKVTERDE 259  
 521 ATIEF-----SFDLERDVPEDAQENVESGVCEGDKTMKAIKDR-- 561

QY	260	DSSINLQKEEFSYFSLRREHDAIQVTHIMAEAGLRISPEQWSSLILYDIAHKSMQSID 319 -VSLLIKRKRQR-QLVRSRKKK-----KBESELLKQOVBQSSAQTGK 604	US-10-052-648-A-40 Sequence 40, Application US/10052648A Publication No. US20040005558A1 GENERAL INFORMATION: APPLICANT: Anderson, David APPLICANT: Burgess, Catherine APPLICANT: Casman, Steven APPLICANT: Colman, Shlomit R. APPLICANT: Edinger, Erik APPLICANT: Elerman, Karen APPLICANT: Gerlach, Valerie APPLICANT: Gunther, Erik APPLICANT: Kekuta, Ramesh APPLICANT: MacDougall, John R. APPLICANT: Mehrabian, Foad APPLICANT: Paturajan, Meera APPLICANT: Rothenberg, Mark APPLICANT: Shimkets, Richard APPLICANT: Smithson, Glenda APPLICANT: Spytek, Kimberly A. APPLICANT: Stone, David J. APPLICANT: Verne, Corine A.M. APPLICANT: Zerhusen, Bryan D.
Db	562	KLQSPEBFASQVQEELTIVLIVGDP-----ANLNRLPHLELLNLDNPDAVSPTW 371	
QY	320	QLPASSTGIFTASTISASVSTQVEEPEPAQDHQQLQYQQPSISVL-----DGTVDSGQSS 662	
Db	605	QLPASSTGIFTASTISASVSTQVEEPEPAQDHQQLQYQQPSISVL-----DGTVDSGQSS 662	
QY	372	EQLENAMVAVKTVVHGLVDNFQTYNSRKGHET-----POQPQNSKYKTSMCRDL 419	
Db	663	VFTESRVSSOCVTVG-----SQEQAHSTGTVPGHIPSTVOAQSPHQGVPPSSVAQG 716	
Db	663	QSQGQ-PSSSLTGYSSQPIQHPQ-----752	
QY	420	RQOGCPCRGNTCTFAHSQEELEKPLRNKINATVRTPFLNK/GVNTTTTAGNIVSV 479	
Db	717	IGSTETTGKIVPSTNGISNAENSQVPSLISRSSTDSTLRALETVKKGKVGANGQNAAGPSA 539	
Db	753	QTVQYSLQ-TSTSSEAT-----TAQPVSQPA 779	
QY	540	DSVTEKIGSPPPKTPVSNVAAATSAGPSN - VGTELNSVPQKSSPFTRV-----PVYPP 591	
Db	780	PQVLP-QVASCKQLPVSSQPVPTIQSBOPIVATQPSVVPVHSGAHLFLPVQCLPPTBLP- 837	
QY	592	HSEN1QYFQDPRTRQFPEVPOY - POTGYVPPPPPTVPG/APCVPFRVRSNNVPBSSLP- 648	
Db	838	QPVSPQIPISTPHVSTAQTGFSSLPLITMAAQTQ-----FLLTASSATAAIPG 887	
QY	649	PASMPYADHYSTSPDRMNNSPY -----OPYP-----POPYGPVPPVPSGMAYPV 694	
Db	888	VSTVVPSQLP-----TLLQPTQQLPSQVHPOULLQPAVQSMGIPANIGQAEBVPLSGDVL 942	
QY	695	YDSRRIWPPMYYQDDIIIRSNSLPPMDVMHSSMVYQTSLSRERYNSLDGYYSVACQPPSEPR 754	
Db	943	YQGFPPRLPQYPGSNIAPSSNVAVSVCIASTVSLPPMPTEVATGFYPTVQVQYVSEN 1002	
QY	755	TTVPL-----PREPGH-----KTSCEBQIRRKPDQWAOYHTOKAPLYSSTLPLVATQ 802	
Db	1003	LLVPPNGGGQVQVQVQVQGQVQVQVQVQV-----TQGVSVQVAPALPVAVA 1054	
QY	803	SP-----TPPSPL-FSYDF-----RADFESVSGTKFEDDLHSHYSPWNSCTIGSCINA 851	
Db	1055	QPOQATOPIITLASSVDAHSDVAGMSDGNENVPSSSGRHEERTKRYHKSVRSSSRHEK 1114	
QY	852	DSEPRDVIAN-----SNAVLMQLDSDGVKRVHLPSTQRRTKEEDPILPFDGPIIISKWG 906	
Db	1115	TSRKLRLRILVNSKGDRTVEOLETHNRKVNTFKFOLDGDNPEELATIMVNNDFLI-----1170	
QY	907	AISRSS-----RIGYHTTDVQ-----TASQ-----SSTATKP 934	
Db	1171	AIEREFSDVQREIIEKADEMLESDEVSPGDQGLESLQGKDGYFGSGSOKLEGFKQP 1230	
QY	935	ISVSDVPPYNAVDSRWSVGNNEATSAHYVERDFITDLSGHKRHS-----TGDL-SL 990	
Db	1231	I PASSMPQIGIPTS-----SLTQVHVSAGR-RFIVSPVPESRRESKVFEPSEITD'TV 1282	
QY	991	ELQQAQSNSLQOREANALAMQKWNLSDEGRHTLNLSKIELRNGELOSDYTEDDATA 1050	
Db	1283	AASTQSPGMNLSSHASSLSLQQAFS-----ELRRAQM-----TEGNT 1321	
QY	1051	TkpDRD-----IEELSLALDDEPDQSEPIEELD-----TOLGISSQNDQ 1092	
Db	1322	APPNTSHGTPPPVPPFLSSIAGVTFRAATAPVATPSSPPNDISTSVQSEVTPTEE 1381	
QY	1093	LNGMAVENGHGPVQHQKEPPQKKOSLGEDHVILEEQTKLIPVTCSESQPLVSYNSAS 1152	
Db	1382	GIAGVATSTGVTSGHPIP-----VSPEPVLSVSSITIPAVVSVSISTS 1428	
QY	1153	-CLPITTS VS 1161	
Db	1429	PSLOVPTSTS 1438	
QY	140	CQIVEEGRVRAARSLGBERTLJLQHONPQOLSANLWAARARGCOFLGPMQEE 1998	
Db	476	HAFFQBETGVRVLAEBDGEKIAKULWRIBDICKLK-----CRYKONE 522	

Qy 200 ALKLVLLAEDGSALSRRKVLVLFVVQRLPFRFQASKTSIGHVQLIYRASCFKVTKRDE 259  
 Db 521 AIEF-----SPDLRDVPDVAQMENVESGTVCEGDHDKTMKAIDR-- 561  
 Qy 260 DSSIMOLKEEFRSYEALRREHDAIQVHIAAMEAGLRISSPEONSSLLYGDLAHKSHMOSIID 319  
 Db 562 -VSLIKRCREQR-QLVREEQERK-----KOEESSLKQVEQSSASOTGK 604  
 Qy 320 KLOSPESPAKSVOBLTIVLQRTGDP-----ANILRRLPHLELLANIDNPNDAVSPTW 371  
 Db 605 QLPASTGJPTASSTTSASVSTQVEPEEPEADQHQQLQXQPSISVLS--DGTVDSCOGSS 662  
 Qy 372 EOLENAMAVAKTVHGLDFIIONYSRKHET-----POPOPNSKYKTSMCRDL 419  
 Db 663 VFTBSRVSQQVTSYG-----SQHEQHSITGTVPHGTPHTVSTVQAQSOPHGVTPPSSVAG 716  
 Qy 420 RQGGCPRGNTNCTFAHSQEELXKYLANKKINATVTRTPFLANKVGVNNTVTTAGNVISV 479  
 Db 717 OSQCQ-PESSSLTGVSSESSQPIQHPQ-----QQQGICQTAPEQ-----752  
 Qy 480 IGSTETTGKIVPSTNGISNAENSQSLISRTSDSTRALEYTKVKGKVANGQNAGPSA 539  
 Db 753 -----QTVQISLQ-TSTSBEAT-----TAQPVSOPQA 779  
 Qy 540 DSYTENKIGSPPKTPVNSAATSAGPSN--VGTELNSVPKQSSPFLTRV-----PVYPP 591  
 Db 780 PQVLP-QVSAGKOLPVSPSVPTIOPQEPOLIPVATQPSVPPVPHSGAHFLPGQPLPTLP- 837  
 Qy 592 HSENIOQYQDPRTOQIPFEPVQY--PQTGYPPPTVPAQAVACVPRFRSNNNVPESSLP- 648  
 Db 838 -----QYPSQIPSTPHYSTAQTGFSLITMAAGITQ--PLTLASSATTAAIPG 887  
 Qy 649 -----PAMPYADHYSTSPPDRMNSSY----QPPP-----PQPKGPVPPVPSGMYAPV 694  
 Db 888 VSTVVPQLP-----TLIQPVTOLPSQVHPQLQPAVQSMGIPANLGQAEEPLSSGCDVL 942  
 Qy 695 YDSRRIMRPPMQRDDIIRSNSLPPDMVHSSVYQTOTSURERYNSLDGYSVACQPPSEPR 754  
 Db 943 YQFPFRPQYQDGSNIAPSNSAVASYVJHSTVLSPPMPEVLTQGYPTVQPYESN 1002  
 Qy 755 TTYPL-----PREPCGHL----KTSCEEQTRRKDQWAOYHTOKAPLYSSTLPVATQ 802  
 Db 1003 LLVBMGGYGGQYQVSOPQSSLAQAPTISQCAVLES-----TQSVQVAPAEPVAVA 1054  
 Qy 803 SP--TPPSPL-FSYDF-----RADSESESGTKEFDLHSYHPSWCGTIGSCINA 851  
 Db 1055 QPQATOPTTLASSVDSAAISDVASGMSDGNEVNPSSSGHEGRUTKRYRSVRSRSRHEK 1114  
 Qy 852 DSEBKDVAN-----SNVALMDLDSGDYKRRVLFETORRTKEBDLIPFSDGPIITSKWG 906  
 Db 1115 TSRPKLRLNNSNKDRVTECQLETHNKKMVTKFEDDGNPEIATIMVNNDPIL--- 1170  
 Qy 907 AISRSS-----RTGYHTTDFOQA-----TASO---GSATKP 934  
 Db 1171 AIERESFYDQVRBIEKADEMLSEDVSYEPEGDOQLESQGKDYGFSGSQKLEGEFKQP 1230  
 Qy 935 ISVSDYVPTVNAVDSRWSYGEATSSHYVERDRFTVTDLSGHRKHS---TGDLL-SL 990  
 Db 1231 IPASSMPQQIGPTS-----SLTQVWHSAGR-RFTVSPVPESRLSRBSKVFPESETIDTV 1282  
 Qy 991 ELQOAKSNLLOREANALAMQKWNSLDEGRHLLNLISKELELRNGELQSDYTEDATD 1050  
 Db 1283 AASTAQSPCMNLSHASSLSQLAFS-----ELRAQM---TEGPNT 1321  
 Qy 1051 TKPDRD-----IELELSALDTDEPDGQSEPIEILD--IOLGISSONDQ 1092  
 Db 1322 APPNFSHSTGPTFPVPPFLSSIAVGPTTAATAAPVPAITSPPNDISTSVIQSSEVTVTEE 1381  
 Qy 1093 LLNGMAVENGHHPVQHQKEPQQKKQUSGEDHVILEEGTILPVTSCSOPULPVSIAMS 1152  
 Db 1382 GIAGVATSTGVVTSGGLPPIP-----VSESPVLSVSVSSTITAVSVISTS 1428

Copyright (c) 1993 - 2005 Compugen Ltd.	GenCore version 5.1.6	Adr15689 Kinase 72	
OM protein - protein search, using SW model		Ads88438 Human pro Adr97079 C. albica	
Run on:	August 4, 2005, 01:25:48 ; Search time 179 Seconds (without alignments)	Aam21438 Peptide # Abb43777 Peptide # Aam37681 Peptide # Abb26719 Peptide #	
Title:	US-10-619-992-2	Aam77509 Human bon Aam64739 Human bra	
Perfect score:	6215 1 MPVQAAQWTBFLSCOPICYNE.....MSEDKNDFLKPVANGKWNNS 1191	Abg59147 Human liv Abg46532 Human pep Abm81216 Tumour-as	
Sequence:	BLOSUM62 Gapop 10.0 , Gapext 0.5	Aab21714 Human WNK Abp71619 Human hea	
Scoring table:	2105692 seqs, 386760381 residues	Adb11119 Human the Adn61859 Human nov Abo44398 Novel hum Abo44382 Novel hum Abo44397 Novel hum	
Searched:	Total number of hits satisfying chosen parameters: 2105692		
ALIGNMENTS			
RESULT 1			
ID:	AAB18608	standard; protein; 1191 AA.	
XX:	AAE18608;		
XX:		AC	
XX:		DT 17-MAY-2002 (first entry)	
DE:	Human cell surface DNA receptor (DNA-R) #1.		
XX:			
KW:	Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;		
KW:	otitis media; septic arthritis; drug screening; gene therapy; virucide;		
KW:	antibacterial; immunosuppressive; chromosome 9q34;		
KW:	auditory.		
XX:			
OS:	Homo sapiens.		
XX:			
PH:		Location/Qualifiers	
FT:		14..50	
FT:		/note= "C3HC3D ring finger domain"	
FT:		122	
FT:		/note= "N-linked glycosylation"	
FT:		394	
FT:		/note= "N-linked glycosylation"	
FT:		416..435	
FT:		/note= "Zinc finger domain"	
FT:		430	
FT:		/note= "N-linked glycosylation"	
FT:		451	
FT:		/note= "N-linked glycosylation"	
FT:		466	
FT:		/note= "N-linked glycosylation"	
FT:		468	
FT:		/note= "N-linked glycosylation"	
FT:		549..809	
Region:		/note= "Proline rich region"	
FT:		1133..1171	
Domain:		/note= "transmembrane domain"	
FT:		1150	
FT:		/note= "N-linked glycosylation"	
XX:			
SUMMARIES			
Result No.:	Score	Query Match %	
		Length DB ID	
- - -	- - -	- - -	
1	6215	100.0 1191 5 AAE18608	Description
2	6207	99.9 1191 5 AAE18609	Aae18608 Human cel
3	6206	99.9 1191 5 AAE18613	Aae18609 Human cel
4	6206	99.9 1191 5 AAE18612	Aae18613 Human cel
5	6205	99.8 1191 5 AAE18610	Aae18612 Human cel
6	5983	88.2 1048 7 ADB80934	Aae18611 Human cel
7	4416	71.1 869 4 AAM40354	Adb80934 RING-SH C
8	4066.5	65.4 812 4 AAM40355	Aam40354 Human pol
9	2665.5	42.9 1125 7 ADC37580	Aam40355 Human pol
10	2577.5	41.5 1109 7 ADF5448	Adc37580 Human nuc
11	2106	33.9 421 4 AAG73504	Adf5448 Human nov
12	1583	25.6 306 4 AAM42140	Aact73504 Human gen
13	1593	25.6 306 4 AAM42141	Aam42140 Human pol
14	1578	24.7 819 4 ABB66041	Aam42141 Human pol
15	1492.5	24.0 296 4 AAG73505	Abb66041 Drosophil
16	611	9.8 522 7 ADB65511	Aag73505 Human gen
17	490	7.9 95 5 ABP64303	Adb65511 Human pro
18	311	5.0 67 5 ABG46802	Abp64303 Human ORF
19	267	4.3 276 4 ABG11134	Abg46802 Human pep
20	205	3.3 1556 8 ADG97964	Abg11134 Novel hum
21	204.5	3.3 127 3 AAB42232	Adg97964 Human can
22	203	3.3 1267 5 ABG65104	Aab42232 Human ORF
23	199.5	3.2 1994 8 ADR18912	Abg65104 Human NOV
24	198.5	3.2 3586 7 ADK40970	Adr18912 Human mac
25			Adk40970 Novel hum

XX	UNIV OREGON HEALTH SCI.	Qy	GSTEITGKIVPSTNGISNAENSVSQIISRSTDSTRALEYTKVKGANGQNAAGPSAD 540
PA	US DEPT VETERANS AFFAIRS.	Db	GSTEITGKIVPSTNGISNAENSVSQIISRSTDSTRALEYTKVKGANGQNAAGPSAD 540
PA	(TARG-) TARGETED GENE DELIVERY.	Qy	SVTENKIGSPPKTPVSNVAATSAGSNVGTELNSVQKSSPFLTRVPPVPHSENQYFQ 600
XX		Db	SVTENKIGSPPKTPVSNVAATSAGSNVGTELNSVQKSSPFLTRVPPVPHSENQYFQ 600
PI	Markins LS, Bennett RM, Seiss DC;	Qy	DPRTOIPPEVPOYPQTGYYPPPTYBAGVAPCVPFRVSNVPPESLPPASMPYADHYST 660
XX		Db	DPRTOIPPEVPOYPQTGYYPPPTYBAGVAPCVPFRVSNVPPESLPPASMPYADHYST 660
WPI:	2002-188747/24.	Qy	FSPRDAMNSSPYQOPPPQPYGPVPPSGMYAPDUDSRR1WPPPMYQDDDLIRNSLPPM 720
DR	N-PSDR, ADD9592.	Db	FSPRDAMNSSPYQOPPPQPYGPVPPSGMYAPDUDSRR1WPPPMYQDDDLIRNSLPPM 720
PS	Page 77-82; 94pp; English.	Qy	DVMHSSVYQTSLRERTNSLDGSYSACQPSPSEPRITVPLPREPCGHLKTSCEBQRKPD 780
XX		Db	DVMHSSVYQTSLRERTNSLDGSYSACQPSPSEPRITVPLPREPCGHLKTSCEBQRKPD 780
CC	The present invention relates to a nucleic acid comprising a sequence encoding a mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA receptor (DNA-R), or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells in vivo and in vitro, and for treating inflammation-associated diseases and conditions, including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogues, and for detecting mammalian DNA-R gene expression in cells and tissues. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express a DNA-R, and the cells are useful as intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R. DNA_R gene is located on chromosome 9q34.	Qy	QWAQHTOKAQLVSTLPLVATOSPTPSPSPLFSVDPRAFDSESVSGTKFEEDHLSHYSPWS 840
CC		Db	QWAQHTOKAQLVSTLPLVATOSPTPSPSPLFSVDPRAFDSESVSGTKFEEDHLSHYSPWS 840
CC		Qy	CGTIGSCINADSEPKDVANSNALMDLGDVKRVRHLFETQRTKEEDPIIPPSDGP 900
CC		Db	CGTIGSCINADSEPKDVANSNALMDLGDVKRVRHLFETQRTKEEDPIIPPSDGP 900
CC		Qy	IISKWGAISRSRTGHTTDPVQATASQSATKPSVSDVPPVNAVDSRMSSYGNBEATS 960
CC		Db	IISKWGAISRSRTGHTTDPVQATASQSATKPSVSDVPPVNAVDSRMSSYGNBEATS 960
CC		Qy	SAHYVERDRFLYTDLSGHRKHSSTGDLSLBLQAKSNSLILQRANALANQKWNSLDE 1020
CC		Db	SAHYVERDRFLYTDLSGHRKHSSTGDLSLBLQAKSNSLILQRANALANQKWNSLDE 1020
CC		Qy	GRHJTLNLSKIEELRNGELOSQYTEDATDKPDRDIELELSALDTDEPDQOSEPTEEL 1080
CC		Db	GRHJTLNLSKIEELRNGELOSQYTEDATDKPDRDIELELSALDTDEPDQOSEPTEEL 1080
CC		Qy	DIOLGISSQNDOLLINGMAVENGHPVQHQKEPPKKQKQSLGEDHVTLEQKTLPPTSCF 1140
CC		Db	DIOLGISSQNDOLLINGMAVENGHPVQHQKEPPKKQKQSLGEDHVTLEQKTLPPTSCF 1140
CC		Qy	SOPLPVYSISNASCCLPITTSSVAGNLLTKTHYMSEIDKNDFLKPVANGRMVNS 1191
CC		Db	SOPLPVYSISNASCCLPITTSSVAGNLLTKTHYMSEIDKNDFLKPVANGRMVNS 1191
SQ	Sequence 1191 AA;	Qy	RESULT 2
Best Local Similarity	100.0%;	Score	6215;
Matches	1191;	DB	5;
Conservative	0;	Pred.	Length 1191;
		No.	0;
		Mismatches	0;
		Indels	0;
		Gaps	0;
Db	1 MPVQAQWTFLSPCIVNEFDENVHKPISLGCSHTVQKTCLINKLHKRACPFQDQTAINTD 60	Qy	1081
Db	1 MPVQAQWTFLSPCIVNEFDENVHKPISLGCSHTVQKTCLINKLHKRACPFQDQTAINTD 60	Db	1081
Qy	61 IDVLPVNFAALLQVGAQWPDHQSIKLNGLGENKHVEYAKKCYVEDLALLYKLPSGGCKGVAS 120	Qy	1141
Db	61 IDVLPVNFAALLQVGAQWPDHQSIKLNGLGENKHVEYAKKCYVEDLALLYKLPSGGCKGVAS 120	Db	1141
Qy	121 LNOSALSRPMQRLVLTLYNCQLVVEEGRVRAARSGLGERVTELLQHNPQQLSANL 180	AAE18609	
Db	121 LNOSALSRPMQRLVLTLYNCQLVVEEGRVRAARSGLGERVTELLQHNPQQLSANL 180	AAE18609	Human cell surface DNA receptor (DNA-R) #2.
Qy	181 WAAVRARGCQFLGPAMQEBALKVLIALEDGSALSRKVLVFLVYORLBRFFQASKTSIG 240	XX	AAE18609 standard; protein; 1191 AA.
Db	181 WAAVRARGCQFLGPAMQEBALKVLIALEDGSALSRKVLVFLVYORLBRFFQASKTSIG 240	XX	
Qy	241 HVVQLLYRASCPFKTVKREDSSMQLKEEFSYEALRREHDQVHIAEGLRISPEQW 300	XX	
Db	241 HVVQLLYRASCPFKTVKREDSSMQLKEEFSYEALRREHDQVHIAEGLRISPEQW 300	XX	Human cell surface DNA receptor; DNA-R; inflammation-associated disease;
Qy	301 SSSLYGDIAHKSHMQS1IDKLQSPESFAKSVQELTIVLQRTD PANLNRPHLEIANI 360	XX	otitis media; septic arthritis; drug screening; gene therapy; virucide;
Db	301 SSSLYGDIAHKSHMQS1IDKLQSPESFAKSVQELTIVLQRTD PANLNRPHLEIANI 360	XX	antiinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
Qy	361 DPNPDAVSPTEQOLENAMAVKTVVHGFLDFTONYSRKGHETPQOPNSKTYKTMCRDLR 420	XX	Homo sapiens.
Db	361 DPNPDAVSPTEQOLENAMAVKTVVHGFLDFTONYSRKGHETPQOPNSKTYKTMCRDLR 420	XX	
Qy	421 QQQGCPCPRTNCTPAHSQEELEYKVRLNKVKINATVTRTPELNVKGVNTVTTAGNVI SVI 480	XX	01-AUG-2001; 2001IWO-US024351.
Db	421 QQQGCPCPRTNCTPAHSQEELEYKVRLNKVKINATVTRTPELNVKGVNTVTTAGNVI SVI 480	XX	01-AUG-2000; 2000US-0222624P.

XX	UNIV OREGON HEALTH SCI.	Qy	481	GSTEITGKIVPSTNGISNAENSVSQIJSRSTDSTLRALEYTVKVGKVGANCONAAGPSAD	540
PA	US DEPT VETERANS AFFAIRS.	Db	481	GSTEITGKIVPSTNGISNAENSVSQIJSRSTDSTLRALEYTVKVGKVGANCONAAGPSAD	540
PA	(TARG-) TARGETED GENE DELIVERY.	Qy	541	SVTENKIGSPPKTPVNSVAATSAGPSMVGTBLNSVPQKSPPFLTRVPPVPPHSENTOYFQ	600
XX		Db	541	SVTENKIGSPPKTPVNSVAATSAGPSMVGTBLNSVPQKSPPFLTRVPPVPPHSENTOYFQ	600
PI	Markins LS, Bennett RM, Seiss DC;	Qy	601	DPTRQIPPEVPPDPPGPPGPPPTPAGVAPCVPRFRSNVNPESLPPASMPYADHYST	660
XX		Db	601	DPTRQIPPEVPPDPPGPPGPPPTPAGVAPCVPRFRSNVNPESLPPASMPYADHYST	660
DR	WPI: 2002-188747/24.	Qy	661	FSPRDRMNSSPYOPPPQGPVPPVPSGMAYAPVYDSRRIWRPPYORDDIRTSNLP	720
DR	N-PSDB, AAD29593.	Db	661	FSPRDRMNSSPYOPPPQGPVPPVPSGMAYAPVYDSRRIWRPPYORDDIRTSNLP	720
XX	New mammalian cell surface DNA receptor proteins and nucleic acids, useful for treating inflammation-associated diseases or conditions, e.g. otitis media, septic arthritis, or bacterial or viral infection that causes inflammation.	Qy	721	DVMHSSYTQTSLERYRNLSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEBQIRRKP	780
PS	Disclosure; Page 88-94; 94pp; English.	Db	721	DVMHSSYTQTSLERYRNLSLDGYSVACQPPSEPRTTVPLPREPCGHLKTSCEBQIRRKP	780
XX	The present invention relates to a nucleic acid comprising a sequence encoding a mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA receptor, or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells in vivo and in vitro, and for treating inflammation-associated diseases and conditions including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogues, and for detecting mammalian DNA-R gene expression in cells and tissues. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express a DNA-R, and the cells are useful as intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R. DNA_R gene is located on chromosome 9q34.	Qy	781	QWAQHTQKAPLVSTLPVATQSTLPVATQSTLPVATQSTLPVATQSTLPVATQSTLPV	840
CC		Db	781	RADFSSESVGSKTFEEDLHSYSPWS	840
CC		Qy	841	CCTIGSCINAIDSEPKDVIANSNAVMLMDLSDGVKCRVHLFETQRRTKEEDPTIPFDGP	900
CC		Db	841	CCTIGSCINAIDSEPKDVIANSNAVMLMDLSDGVKCRVHLFETQRRTKEEDPTIPFDGP	900
CC		Qy	901	IISKGWAISSRSRTGYHTDPYQATASQGSAATKPIISVSDYYPVYNAVDSRNNSYGNBEAT	960
CC		Db	901	IISKGWAISSRSRTGYHTDPYQATASQGSAATKPIISVSDYYPVYNAVDSRNNSYGNBEAT	960
CC		Qy	961	SAHYVERDRFIVTDLSGHRKHSSTGDSLUSLLEQAKNSNLQREANALAMQKWNSLDB	1020
CC		Db	961	SAHYVERDRFIVTDLSGHRKHSSTGDSLUSLLEQAKNSNLQREANALAMQKWNSLDB	1020
XX		Qy	1021	GRHLTLNLSSKTELRNGELQSDYTEDATDKPDRDTELESALDTDPDGQSEPIEBIL	1080
SQ	Sequence 1191 AA:	Db	1021	GRHLTLNLSSKTELRNGELQSDYTEDATDKPDRDTELESALDTDPDGQSEPIEBIL	1080
Query Match	99.9% ; Score 6207; DB 5; Length 1191;	Qy	1081	DIQGLISSQNDOLLNGMAVENCHPVQOHQKEPKQKKQSLGEDDHVILEQKTILPVTS	1140
Best Local Similarity	99.9%; Pred. No. 0; Gaps 0;	Db	1081	DIQGLISSQNDOLLNGMAVENCHPVQOHQKEPKQKKQSLGEDDHVILEQKTILPVTS	1140
Matches	1190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	Qy	1141	SOPLPVLSISNASCLPITTSVAGNLILKTHYNSEDKNDPLKPVANGRPNVS	1191
Qy	1 MPVOAAQTTEPLSPCIPYNEFDENVHKPISLGCSHTYCKTCLNKLRKACCPFDQTAINTD 60	Db	1141	SOPLPVLSISNASCLPITTSVAGNLILKTHYNSEDKNDPLKPVANGRPNVS	1191
Db	1 MAYOOAQNTTEFLSPCIPYNEFDENVHKPISLGCSHTYCKTCLNKLRKACCPFDQTAINTD 60	Qy	1141	SOPLPVLSISNASCLPITTSVAGNLILKTHYNSEDKNDPLKPVANGRPNVS	1191
Qy	61 IDVLPVNFAALLQVGAQPDHOSTKLNLGENGHYEV/KKCYVEDDLALYTKPLSGGGKCVAS 120	Db	1141	SOPLPVLSISNASCLPITTSVAGNLILKTHYNSEDKNDPLKPVANGRPNVS	1191
Db	61 IDVLPVNFAALLQVGAQPDHOSTKLNLGENGHYEV/KKCYVEDDLALYTKPLSGGGKCVAS 120	Qy	121 LNQSLALSERMQRKLTVLNQLVLEEGVRAARASIGERTVTELLQHNPQQLSANL 180	RESULT 3	
Db	121 LNQSLALSERMQRKLTVLNQLVLEEGVRAARASIGERTVTELLQHNPQQLSANL 180	ID	AAB18613 standard; protein; 1191 AA.		
Qy	181 WAAVRARGCQFLGPAMOEEALKLVLALLEDGSALSRKVYVLFVQRLPFPQASKTSIG 240	XX			
Db	181 WAAVRARGCQFLGPAMOEEALKLVLALLEDGSALSRKVYVLFVQRLPFPQASKTSIG 240	AC	AAB18613;		
Qy	241 HVVQLLYRASCFCYTKTKEDESSLQMKKEFRSTEARLREHDAQVH1AMEAGHRISPEQW 300	XX			
Db	241 HVVQLLYRASCFCYTKTKEDESSLQMKKEFRSTEARLREHDAQVH1AMEAGHRISPEQW 300	KW	Human; cell surface DNA receptor; DNA-R; inflammation-associated disease;		
Qy	301 SLLIYGDLAHKSHMQSIIDKLQSPESFAKSQVQELTIVQRTGPANJNLRPHLELLANI 360	KW	oritis media; septic arthritis; drug screening; Gene therapy; virucide;		
Db	301 SLLIYGDLAHKSHMQSIIDKLQSPESFAKSQVQELTIVQRTGPANJNLRPHLELLANI 360	KW	antiinflammatory; antibacterial; immunosuppressive; Chromosome 9q34;		
Qy	361 DPNDPAVSPTEQWELENAYAVKTVYGLDFIONYSRKGHEHTPOPQNSKYSKTSMCRDLR 420	KW	auditory; mutant; mutain.		
Db	361 DPNDPAVSPTEQWELENAYAVKTVYGLDFIONYSRKGHEHTPOPQNSKYSKTSMCRDLR 420	KW	Homo sapiens.		
Qy	421 QGCGCPRGTMCTPAHSQEELEYKYLNRKINNATRTEFLINKYGVANTVTTAGNVNISI 480	XX			
Db	421 QGCGCPRGTMCTPAHSQEELEYKYLNRKINNATRTEFLINKYGVANTVTTAGNVNISI 480	FT	Location/Qualifiers 431		
Qy	421 QGCGCPRGTMCTPAHSQEELEYKYLNRKINNATRTEFLINKYGVANTVTTAGNVNISI 480	FT	/note= "Wild type Cys substituted with Ala"		
Db	421 QGCGCPRGTMCTPAHSQEELEYKYLNRKINNATRTEFLINKYGVANTVTTAGNVNISI 480	XX	WO200210392-A2.		
PD		XX	PD 07-FEB-2002.		

XX	01-AUG-2001; 2001WO-US024351.	Db	361 DPNPDAVSPITWQLELNAMEAVAKTVVHGLVDEIQNYSRKGHETPQQQPNSKYKTSMCRDLR 420
PF	PR 01-AUG-2000; 2000US -0222624P.	Qy	421 QGGCPRGTNTCTFAHSQEELFKYRLRNKINATVTFPLINKVG/NNTVT/TTAGNVISVI 480
XX	UNIV OREGON HEALTH SCI	Db	421 QGGCPRGTNTCTFAHSQEELFKYRLRNKINATVTFPLINKVG/NNTVT/TTAGNVISVI 480
PA	(USGO ) US DEPT VETERANS AFFAIRS.	Qy	481 GSTETGKIVPSTNGISNAENSQQLISRSTDSTALETYKKVCGANGQNAAGPSAD 540
PA	(TARG-) TARGETED GENE DELIVERY.	Db	481 GSTETGKIVPSTNGISNAENSQQLISRSTDSTALETYKKVCGANGQNAAGPSAD 540
XX	PI Hefeneider S, Merkins LS, Bennett RM, Seiss DC;	Qy	541 SVTENKIGSPKTPVNSVAATSAGGPNGVTELNSPKQSSPFLTVTPVTPPHSENQYFQ 600
XX	DR 2002-188747/24.	Db	541 SVTENKIGSPKTPVNSVAATSAGGSNGVTELNSPKQSSPFLTVTPVTPPHSENQYFQ 600
PT	New mammalian cell surface DNA receptor proteins and nucleic acids, useful for treating inflammation-associated diseases or conditions, e.g. otitis media, septic arthritis, or bacterial or viral infection that causes inflammation.	Qy	601 DPRTQIPFEVQPOYPOTGYYPPRPTVAGVACPVPEFRVSNNVPESLPPASMPYADHYST 660
PT	PT XX	Db	601 DPRTQIPFEVQPOYPOTGYYPPRPTVAGVACPVPEFRVSNNVPESLPPASMPYADHYST 660
PS	PS Example 2; Page: 94pp; English.	Qy	661 FSPRDRMNSSSPYQPPRQPYGPVPPSPSGMYAPVYDSDRRIWRRPMYQRDDDTIRSNLSLPPM 720
XX	The present invention relates to a nucleic acid comprising a sequence encoding a mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells in vivo and in vitro, and for treating inflammation-associated diseases and conditions including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogues, and for detecting mammalian DNA-R gene expression in cells and tissues. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express a DNA-R, and the cells are useful as intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R mutant. DNA-R gene is located on chromosome 9q34. Note: The present sequence is not shown in the specification, but is derived from the wild type DNA-R sequence shown in sequence listing (page 77-82) of the specification.	Db	661 FSPRDRMNSSSPYQPPRQPYGPVPPSPSGMYAPVYDSDRRIWRRPMYQRDDDTIRSNLSLPPM 720
CC	CC CC	Qy	721 DMHSSVYQTSLRERYNSLDGYYSAQPPBEPRTTVPLPREPCGHLKTSCEBQRRKPD 780
CC	CC CC	Db	721 DMHSSVYQTSLRERYNSLDGYYSAQPPBEPRTTVPLPREPCGHLKTSCEBQRRKPD 780
CC	CC CC	Qy	781 QWAQHTOKAPLVSSLPLVATOSPTPPSPLESVDPRADFSESVSGTKFEDDHLSHYSPWS 840
CC	CC CC	Db	781 QWAQHTQTQAKLVLSTLPVATOSPTPPSPLESVDPRADFSESVSGTKFEDDHLSHYSPWS 840
CC	CC CC	Qy	841 CGTIGSCINADSEPKDVANSNATVMLDLSGDGVKRVRHLFETQRTKEEDPPIPPSDGP 900
CC	CC CC	Db	841 CGTIGSCINADSEPKDVANSNATVMLDLSGDGVKRVRHLFETQRTKEEDPPIPPSDGP 900
CC	CC CC	Qy	901 IISKWGAISRSRTGHTTDPVQATASQGATKPSVSDYVPPVNAVDWSMSSYNEATS 960
CC	CC CC	Db	901 IISKWGAISRSRTGHTTDPVQATASQGATKPSVSDYVPPVNAVDWSMSSYNEATS 960
CC	CC CC	Qy	961 SAHYVERDRFIVTDLSGHRKHSSTCDPLSLLEQOKRSNSLLQRANALAMQKQNSLDE 1020
CC	CC CC	Db	961 SAHYVERDRFIVTDLSGHRKHSSTCDPLSLLEQOKRSNSLLQRANALAMQKQNSLDE 1020
SQ	Sequence 1191 AA;	Qy	1021 GRHLTTLNSKELIERNGLQSDYTEDATDKPDRDIELELSALDTDEPDQOSEPTEEIL 1080
XX	Query Match 99.9%; Score 6206; DB 5; Length 1191; Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0; Matches 1190; Conservative 0; Gaps 0;	Db	1021 GRHLTTLNSKELIERNGLQSDYTEDATDKPDRDIELELSALDTDEPDQOSEPTEEIL 1080
Qy	1 MPVQAAQWTEFLSPCIPYNEFDENVHKPISLGCSHTVCKTCCLNKHLRKACPFDTAINTD 60	Qy	1081 DIQLGISSQNQDOLLNGMAVENGHPOQHQKEPKKKQKQSLGEDDHYTLEQKTIILPTVTSFC 1140
Db	1 MPVQAAQWTEFLSPCIPYNEFDENVHKPISLGCSHTVCKTCCLNKHLRKACPFDTAINTD 60	Db	1081 DIQLGISSQNQDOLLNGMAVENGHPOQHQKEPKKKQKQSLGEDDHYTLEQKTIILPTVTSFC 1140
Qy	61 IDVLPIFNFAOLQLVGQVDPDQSITKLSNLGENKHYEVAKCVCEDDLALYKLPLSGGGKVVAS 120	Qy	1141 SQPLPVSIASNACLPIITTSAGNILLKTHVMSEDFNDFLKPVANGKVMNS 1191
Db	61 IDVLPIFNFAOLQLVGQVDPDQSITKLSNLGENKHYEVAKCVCEDDLALYKLPLSGGGKVVAS 120	Db	1141 SQPLPVSIASNACLPIITTSAGNILLKTHVMSEDFNDFLKPVANGKVMNS 1191
Qy	121 LNQSA1SRPMQRLVTLVNCOLVEEERVRAARSLGERVTTELLQHNPQQLSANL 180	RESULT 4	
Db	121 LNQSA1SRPMQRLVTLVNCOLVEEERVRAARSLGERVTTELLQHNPQQLSANL 180	AAE18612	
Qy	181 WAAVRARGCQFLGPMQMOBEALKLVLIALEDGSALSRSKVLVLFVORLEPRFQASKTSIG 240	ID	AAE18612 standard; protein; 1191 AA.
Db	181 WAAVRARGCQFLGPMQMOBEALKLVLIALEDGSALSRSKVLVLFVORLEPRFQASKTSIG 240	XX	
Qy	241 HVQQLYXASCFKVTRDEDSSMOLKEEFSYEALRREHDQVHIAAMEGLRISPQEW 300	AC	
Db	241 HVQQLYXASCFKVTRDEDSSMOLKEEFSYEALRREHDQVHIAAMEGLRISPQEW 300	XX	
Qy	301 SSILYGDIAHKSMSIDIKLQSPESAKSVOBLTVLQRTGDPANLRPHLELLANI 360	DT	17-MAY-2002 (first entry)
Db	301 SSILYGDIAHKSMSIDIKLQSPESAKSVOBLTVLQRTGDPANLRPHLELLANI 360	XX	
Qy	361 DPNPDAVSPITWQLELNAMEAVAKTVVHGLVDEIQNYSRKGHETPQQQPNSKYKTSMCRDLR 420	OS	
		XX	
		Key	Location/Qualifiers

FT	Misc-difference 416 /note= "Wild type Cys substituted with Ala"	Qy	301 SSLIYGDLAHKSHMMQSTIDKLQSPESPAKSVOBLTTIVLQRTGDPANLNRLPHLELLANI 360
FT	XX	Db	301 SSLIYGDLAHKSHMMQSTIDKLQSPESPAKSVOBLTTIVLQRTGDPANLNRLPHLELLANI 360
EN	W0200210392-A2.	Qy	361 DPNPDAVSPTEQLENAMAVAKTVVGLVDFQNSRKGHETPQQPQDNSSKKTSMCBDLR 420
XX	07-FEB-2002.	Db	361 DPNPDAVSPTEQLENAMAVAKTVVGLVDFQNSRKGHETPQQPQDNSSKKTSMARDLR 420
XX	01-AUG-2001; 2001WO-US024351.	Qy	421 QGGCPRGTCNTFAHSDELEYRLRANKKINATVTRTPLLNKVGVNNTVTITAGNIVSVI 480
PR	01-AUG-2000; 2000US-0222624P.	Db	421 QGGCPGTGTCNTFAHSDELEYRLRANKKINATVTRTPLLNKVGVNNTVTITAGNIVSVI 480
XX	(UYOR - UNIV OREGON HEALTH SCI.	Qy	481 GSTETTGSKIVPSTNGISNAENSVSQIISRSTDSTRALEYTKVKGANGONAGPSAD 540
PA	(USGO ) US DEPT VETERANS AFFAIRS.	Db	481 GSTETTGSKIVPSTNGISNAENSVSQIISRSTDSTRALEYTKVKGANGONAGPSAD 540
PA	(TARG-) TARGETED GENE DELIVERY.	Qy	
XX	Hefeneider S, Merkins LS, Bennett RM, Seiss DC;	Db	
PI	WPI; 2002-188747/24.	Qy	541 SYTENKIGSPPKTPVSNVAATSAGSPNGTELNSPQKSSPPLTRVPPVPPHSENQYFQ 600
XX	DR;	Db	541 SYTENKIGSPPKTPVSNVAATSAGSPNGTELNSPQKSSPPLTRVPPVPPHSENQYFQ 600
PR	New mammalian cell surface DNA receptor proteins and nucleic acids, or nucleic acid sequences associated with diseases or conditions, e.g., useful for treating inflammation-associated diseases or conditions, e.g., otitis media, septic arthritis, or bacterial or viral infection that causes inflammation.	Qy	601 DPTQIIPFEVPOVQPTQYPPPPTVPGVAPCPVFRTSNNVPESSLPPASHYPTDHYST 660
PR	PR	Db	601 DPTQIIPFEVPOVQPTQYPPPPTVPGVAPCPVFRTSNNVPESSLPPASHYPTDHYST 660
PS	Example 2; Page; 94pp; English.	Qy	661 FSPRDRKNSSPYOPPPQPYGPVPPSGMAYPAVDSSRRIWRPPMYYDODDIIRNSLPPM 720
XX	DR;	Db	661 FSPRDRKNSSPYOPPPQPYGPVPPSGMAYPAVDSSRRIWRPPMYYDODDIIRNSLPPM 720
CC	The present invention relates to a nucleic acid comprising a sequence encoding mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells in vivo and in vitro, and for treating inflammation-associated diseases and conditions including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogues, and for detecting mammalian DNA-R gene expression in cells and tissues. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express a DNA-R, and the cells are useful as intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R mutant. DNA-R gene is located on chromosome 9q34. Note: The present sequence is not shown in the specification, but is derived from the wild type DNA-R sequence shown in sequence listing (page 77-82) of the specification.	Qy	721 DYMHSYYQTSLRERYSNLGDGYSVACOPPSERPTVPLPREPCGHULKTSCEBQIRRKPD 780
CC	XX	Db	721 DYMHSYYQTSLRERYSNLGDGYSVACOPPSERPTVPLPREPCGHULKTSCEBQIRRKPD 780
CC	PS	Qy	781 QHAQHHTQKAPLYSSTLIPVATOSPTPPSPLFSDYFRADFSESVGTFKEDHLSHYSWMS 840
CC	Example 2; Page; 94pp; English.	Db	781 QHAQHHTQKAPLYSSTLIPVATOSPTPPSPLFSDYFRADFSESVGTFKEDHLSHYSWMS 840
CC	XX	Qy	841 CGTIGSCINATDSEPKDVANSVAVMDLSDGPKVKEVHLFETQRRTKEEDPPIPFSGP 900
CC	DR;	Db	841 CGTIGSCINATDSEPKDVANSVAVMDLSDGPKVKEVHLFETQRRTKEEDPPIPFSGP 900
CC	PR	Qy	901 IISKGWALRSRSSRTGYHTTDPTVQATASOGSATKPKISVSDVYVNAVDSRWSSYGNBEAT 960
CC	PS	Db	901 IISKGWALRSRSSRTGYHTTDPTVQATASOGSATKPKISVSDVYVNAVDSRWSSYGNBEAT 960
CC	XX	Qy	961 SAHYVERDRFIVTDSLGRKHGSTGDLISLELQQAKNSNLLQREANALAMQKWNSLDE 1020
CC	DR;	Db	961 SAHYVERDRFIVTDSLGRKHGSTGDLISLELQQAKNSNLLQREANALAMQKWNSLDE 1020
SQ	Sequence 1191 AA;	Qy	1021 GRHLTLNLSKEELRNCELGQSDTYTEDATDKPDRDIELELSALDTDEPDGQSEPIEIL 1080
CC	Best Local Similarity 99.9%; Pred. No. 0; Mismatches 1; Indels 0; Gaps 0;	Db	1021 GRHLTLNLSKEELRNCELGQSDTYTEDATDKPDRDIELELSALDTDEPDGQSEPIEIL 1080
CC	Matches 1190; Conservative	Qy	1081 DIOGLISSQNDQQLNGMAVENGHGPVQOQKEPPKKQKQSLGEDPHVILEEQKTLIPVTSFC 1140
DB	1 MPVOAAQTBFSLCPICNEFDENVHPKTSLGCSHTVCKTCLNKLHRACPFQDQTAINTD 60	Db	1081 DIOGLISSQNDQQLNGMAVENGHGPVQOQKEPPKKQKQSLGEDPHVILEEQKTLIPVTSFC 1140
DB	1 MPVOAAQTBFSLCPICNEFDENVHPKTSLGCSHTVCKTCLNKLHRACPFQDQTAINTD 60	Qy	1141 SQPLPVSIISNASCLPITTSVAGNLILKTHNSEDRDFLKPVANGKVNNS 1191
QY	61 IDVLPPNVNFAALLQVGAQPDHOSIKLSNLGENGHYEVAVKCVDLTLKPLSGGKVAS 120	Db	1141 SQPLPVSIISNASCLPITTSVAGNLILKTHNSEDRDFLKPVANGKVNNS 1191
DB	61 IDVLPPNVNFAALLQVGAQPDHOSIKLSNLGENGHYEVAVKCVDLTLKPLSGGKVAS 120	Qy	
QY	121 LNQNALSRPMORKLUVTLNCOLVVEEGTRMARAARSGERTYTELLIOHQNPQLSNL 180	Db	
DB	121 LNQNALSRPMORKLUVTLNCOLVVEEGTRMARAARSGERTYTELLIOHQNPQLSNL 180	Qy	
QY	1B1 WAAVRARGCQFLGPAMQEEALKLVLAEDGSAALSRSKTVLFLVQRLPREFPQASCTSIG 240	Db	
DB	1B1 WAAVRARGCQFLGPAMQEEALKLVLAEDGSAALSRSKTVLFLVQRLPREFPQASCTSIG 240	Qy	
QY	241 HVVQQLYRASCFCYKVKRDDESSLMQLKEEPRSTEARREHDAQJVHIAEAGLRISPPSQW 300	Db	
DB	241 HVVQQLYRASCFCYKVKRDDESSLMQLKEEPRSTEARREHDAQJVHIAEAGLRISPPSQW 300	Qy	
XX	Human cell surface DNA receptor (DNA-R) #1 mutant, C431S.	XX	
KW	Human; cell surface DNA receptor; DNA-R; inflammation-associated disease; otitis media; septic arthritis; drug screening; gene therapy; viricide;	KW	

KRW	antinflammatory; antibacterial; immunosuppressive; chromosome 9q34;
KNW	auditory; mutant; mutein.
XXX	
QOS	
XX	
Homo sapiens.	
	Key Location/Qualifiers
	Misc-difference 431 /note= "Wild type Cys substituted with Ser"
W0200210392-A2.	
PN	
PN	
XXX	
IPD	07-FEB-2002.
XXX	
PPF	01-AUG-2001; 2001WO-US024351.
XX	
PR	01-AUG-2000; 2000US-0222624P.
XXX	
PA	(UTOR-) UNIV OREGON HEALTH SCI.
PA	(USGO ) US DEPT VETERANS AFFAIRS.
PA	(TARG-) TARGETED GENE DELIVERY.
XX	
PPI	HeFeneider S, Merkins LS, Bennett RM, Seiss DC;
XX	
DR	WPI: 2002-188747/24.
XXX	
PPT	New mammalian cell surface DNA receptor proteins and nucleic acids, e
PPT	useful for treating inflammation-associated diseases or conditions, e
PPT	otitis media, septic arthritis, or bacterial or viral infection that
PPT	causes inflammation.
PPT	

Example 2: Page: 94/p; English.

The present invention relates to a nucleic acid comprising a sequence encoding a mammalian cell surface DNA receptor (DNA-R), a soluble mammalian cell surface DNA-R, or a DNA-binding fragment of a mammalian DNA-R. Mammalian DNA-R proteins may be used to screen compounds that effect DNA binding to cells *in vivo* and *in vitro*, and for treating inflammation-associated diseases and conditions including otitis media, septic arthritis, and any bacterial or viral infection that causes inflammation by interaction with the DNA-R. These may also be used to screen compounds that modulate binding, uptake and expression. The nucleic acid probes are useful for isolating mammalian species analogs and for detecting mammalian DNA-R gene expression in cells and tissue. Recombinant expression constructs are useful in molecular biology to transform cells which do not ordinarily express DNA-R, and the cell intermediates for making cell membrane preparations for receptor binding assays, which are subsequently useful in drug screening. The recombinant constructs are also useful in gene therapy. The present sequence is human DNA-R mutant. DNA-R Gene is located on chromosome 9. Note: The present sequence is not shown in the specification but is derived from the wild type DNA-R sequence shown in sequence listing (77-82) of the specification.

Sequence 1191 AA:			
Seq	Query Match	Score 6205;	Length 1191;
	Best Local Similarity	99.8%; 99.9%;	Prod. No. 0; T-221- 0 00000

1 MPVOAAOWTEELSCRPIYNEEDENVHKRISLGCSHTYSETCLNKLHRKACPFDOTAINTR  
Matches 1190; Conservative 0; Mismatches 1; Indels 0; gaps 0

1 MPVQAAQWTEFLSPICYNEFEDENVKPISLGCSHTVKTCNLKHLRKACPFDQTANTD

61 DVLPVNFAALLQVGAQVPDPHOSIKLNSLGENKHYEVAKKCVEDLALYLKPLSGGKCVASQY

61 IDVLPVNPALLQLVGAQVPDHOISIKLNSNLGENKTYEVAKKCVEDALYLKPSSGGKTYAS

	Qy	Db	Qy	
121	INQ SAL SRP MQR KLV TLV NC QL VEE EGR VRA AAR SUGERT TEL LQH NP QOLS ANI	INQ SAL SRP MQR KLV TLV NC QL VEE EGR VRA AAR SUGERT TEL LQH NP QOLS ANI	121	INQ SAL SRP MQR KLV TLV NC QL VEE EGR VRA AAR SUGERT TEL LQH NP QOLS ANI
181	WAA VR A R G C Q F G P A M Q F E A K U V I L L A L E D G S A L S R K V I V L V F Y O R L E P R F P Q A S K T S G			

181	WAAVRARGQFLGAMQEAKLKVIALLEGSAALKVLFVQQBLEPRFQASKTSIG	240
241	HVQQLYRASCFKTYKTRDEDSLMLKEEPRSYEARREHDAIQVHTIAEGLRISPEQW	300
241	HVQQLYRASCFKTYKTRDEDSLMLKEEPRSYEARREHDAIQVHTIAEGLRISPEQW	300
301	SSLLYGLAHKSHMCSIDKLQSPSPAKSVOELTIVLQRTGDPANLNRPHLELLANI	360
301	SSLLYGLAHKSHMCSIDKLQSPSPAKSVOELTIVLQRTGDPANLNRPHLELLANI	360
361	DPNPDAVSPTWEQLENAMYAVKTIVHGLYDFTQNSRSRKGETPOPOENSKYKTSMCRLR	420
361	DPNPDAVSPTWEQLENAMYAVKTIVHGLYDFTQNSRSRKGETPOPOENSKYKTSMCRLR	420
421	QOGCOPRGNTCTPAHQEELEKYLRLRKCKINATVTPPLANKVGUNNTVTITTAGNVTSVI	480
421	QOGCOPRGNTCTPAHQEELEKYLRLRKCKINATVTPPLANKVGUNNTVTITTAGNVTSVI	480
481	GSTETTGK1VPSTNGISNAENSQSLI1SRSTDSTRALEYTKVKVGANGONAAGPBD	540
481	GSTETTGK1VPSTNGISNAENSQSLI1SRSTDSTRALEYTKVKVGANGONAAGPBD	540
541	SYTERNKIGSPPKTPVSNVAATSAGPSNVGTTELNSYPKQKSSPFLTRVYPVPPHSENIOYFO	600
541	SYTERNKIGSPPKTPVSNVAATSAGPSNVGTTELNSYPKQKSSPFLTRVYPVPPHSENIOYFO	600
601	DPRTOQIPPEVPOYPOQTGYYPPPPTVAGAPCVPFRVSNVNPESLPPASMPYADHYST	660
601	DPRTOQIPPEVPOYPOQTGYYPPPPTVAGAPCVPFRVSNVNPESLPPASMPYADHYST	660
661	FSPRDRMNSPYQPQPPQPYGPVPPVPSGMAYAPVDSRR1WRPMYORDDIRNSLPPM	720
661	FSPRDRMNSPYQPQPPQPYGPVPPVPSGMAYAPVDSRR1WRPMYORDDIRNSLPPM	720
721	DVMHSIVYQTSLRREBYNSLUDGGSYVACQPPSEPRTRYPLPREPCGHLKTSCEQIRRKPD	780
721	DVMHSIVYQTSLRREBYNSLUDGGSYVACQPPSEPRTRYPLPREPCGHLKTSCEQIRRKPD	780
781	QWAQHTOKAPLVSTLPLVATQSPTPPSPLFSVDFADESESVSGTKPFEEDHLSHYSPWS	840
781	QWAQHTOKAPLVSTLPLVATQSPTPPSPLFSVDFADESESVS GTKPFEEDHLSHYSPWS	840
841	CGTICSCINADSEPKDVIANSAVIMLDLQSGDVRRVHLFBTQRTRKEEDP1PPFSDGP	900
841	CGTICSCINADSEPKDVIANSAVIMLDLQSGDVRRVHLFBTQRTRKEEDP1PPFSDGP	900
901	IISKGAISSRSSRTGHTTOPVQATASQGSAATKPLSVSDYYVNAVDSRWSSYNEATS	960
901	IISKGAISSRSSRTGHTTOPVQATASQGSAATKPLSVSDYYVNAVDSRWSSYNEATS	960
961	SAHYTERDR1VTDLGHRKHSSTGDLISLQAKNSNLLIQRANALAMQKWNLSDE	102
961	SAHYTERDR1VTDLGHRKHSSTGDLISLQAKNSNLLIQRANALAMQKWNLSDE	102
1021	GRHLTLNLISKEIELRNGELOSQDTEDATDKPDRDTELESALDTDPDGQBEPBIEEL	108
1021	GRHLTLNLISKEIELRNGELOSQDTEDATDKPDRDTELESALDTDPDGQBEPBIEEL	108
1081	DIOLGISSONDQ1LNGMAVENGHPIQOHOKEPPKKQKOSLGEDDHVILEEQKTLPPVTSFC	114
1081	DIOLGISSONDQ1LNGMAVENGHPIQOHOKEPPKKQKOSLGEDDHVILEEQKTLPPVTSFC	114
1141	SQPLPVSIENASCLPITTSAGNLLKTHMSEDKNDLKEPVANGKVNNS	1191
1141	SQPLPVSIENASCLPITTSAGNLLKTHMSEDKNDLKEPVANGKVNNS	1191



RESULT 7							
ADB80934	standard; protein, 1048 AA.						
XX							
AC	ADB80934;						
XX							
DT	04-DEC-2003 (first entry)						
XX							
DE	RING-SH complex related protein, SEQ ID No 8.						
XX							
RING-SH 3; Gag protein; Gag late domain; PI3K; actin; myosin; Hsp60; Hep70; Hsp90; STAM1; STAM2A; VHS-UIM; GRPase; E2 enzyme; tsg101; cullin; RING-SH; clathrin; virucide; vaccine; antiviral; retrovirus; rhabdovirus; filovirus.							
XX	Unidentified.						
XX	W02003033646-A2.						
XX	PD 24-APR-2003.						
XX	PP 31-JUL-2002; 2004WO-US024559.						
XX	PR 31-JUL-2001; 2001US-0308958P.						
PR 09-NOV-2001; 2001US-0345846P.							
XX	(PROT-) PROTEOLOGICS INC.						
XX	PR Greener T, Moskowitz H, Reiss Y, Alroy I;						
XX	DR WPI; 2003-293509/37.						
DR N-PSDB; ADB80975.							
XX	New isolated protein complex comprising a RING-SH 3 polypeptide and another polypeptide, useful for detecting cells infected with a virus, and for treating viral disorders caused by retroviruses, rhabdoviruses, or filoviruses.						
XX	Disclosure; Fig 8; 176pp; English.						
CC	The invention relates to a novel isolated protein complex comprising a RING-SH 3 polypeptide and a polypeptide selected from a Gag protein, a Gag late domain, PI3K, actin, myosin, Hsp60, Hsp70, Hsp90, STAM1, STAM2A, STAM2B, VHS-UIM, a Grpase, an E2 enzyme, tsg101, a cullin, RING-SH, and a clathrin. The novel protein complex has virucide activity and can be used to treat disorders as part of a vaccine. The protein complex and its composition are useful for detecting cells infected with a virus, for identifying agents having antiviral activity and for treating viral disorders caused by retroviruses, rhabdoviruses, or filoviruses. This sequence is a protein comprising the RING-SH complex of the invention.						
PS	Sequence 1048 AA.						
QQ	Query Match 88.2%; Score 5483; DB 7; Length 1048; Best Local Similarity 100.0%; Pred. No. 0; Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Db	1 MPYQAAQTEFLSCPICNEFDENVHKIEFISLGCSHTIVCTCLNKLARKACPPDQTANTD 1 MPYQAAQTEFLSCPICNEFDENVHKIEFISLGCSHTIVCTCLNKLARKACPPDQTANTD	60					
Db	61 IDVLPVNFAALLQVGAQYDHQSISIKLSNLGENHRYAKCVDLALYLPUSGGRCVAS 61 IDVLPVNFAALLQVGAQYDHQSISIKLSNLGENHRYAKCVDLALYLPUSGGRCVAS	120					
Db	61 LNQSLASPRMQLKLVTLVNCOLVYEEGRVRAARSIGERTYTELIIQHNPQQLSANT 61 LNQSLASPRMQLKLVTLVNCOLVYEEGRVRAARSIGERTYTELIIQHNPQQLSANT	180					
Qy	121 LNQSLASPRMQLKLVTLVNCOLVYEEGRVRAARSIGERTYTELIIQHNPQQLSANT 121 LNQSLASPRMQLKLVTLVNCOLVYEEGRVRAARSIGERTYTELIIQHNPQQLSANT	180					
Db	181 WAAVRARGCQFLGPMQSEALKLVLAEDGSLSRCKTUVLVRPRFOQASITSIG 181 WAAVRARGCQFLGPMQSEALKLVLAEDGSLSRCKTUVLVRPRFOQASITSIG	240					
Qy	181 WAAVRARGCQFLGPMQSEALKLVLAEDGSLSRCKTUVLVRPRFOQASITSIG 181 WAAVRARGCQFLGPMQSEALKLVLAEDGSLSRCKTUVLVRPRFOQASITSIG	240					
Db							
RESULT 8							
ID	AAM40354 standard; protein; 869 AA.						
XX							
AC	AAM40354;						
XX	22-OCT-2001 (First entry)						
DE	Human polypeptide SEQ ID NO 3499.						
XX							
KW	Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chmotoxic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.						



PR	03-AUG-2000; 2000US-00653450.	Qy	676 PPQPGPVPVPSGMATAPVYDVSRRIRPMPYQDDIIRNSLPPMDVMHSSVYQTSRLR 735
PR	14-SEP-2000; 2000US-00662391.	Db	481 PPQPGPVPVPSGMATAPVYDVSRRIRPMPYQDDIIRNSLPPMDVMHSSVYQTSRLR 540
PR	19-OCT-2000; 2000US-00693346.	Qy	736 YNSLDGYSVACOPSEPRTPVPLPREPCSHLKTCSEBQIRRKPDQWAQHTQKPLVSS 795
PR	29-NOV-2000; 2000US-00727344.	Db	541 YNSLDGYSVACOPSEPRTPVPLR----- 566
XX	(HYSEQ INC.	Qy	796 TLPVATOSPTPPSPLESVDFPRADPESESQTKFEDHLHYSPNCGTIGSCINADSEP 855
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	Db	567 ----- -FEBSVSGTKEFDHLSHYSXPWCGTIGSCINADSEP 603
PI	Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	Qy	856 KDVIVANSNAVLMQDLSGDVGRVHFETQRTRKEEDPPIPSDGPDISKMGAISSSRTG 915
PI	Zhou P, Goodrich R, Drmanac RT;	Db	604 KDVIVANSNAVLMQDLSGDVGRVHFETQRTRKEEDPPIPSDGPDISKMGAISSSRTG 663
XX	DR; 2001-442253/47.	Qy	916 YHTTDPVQATASQGATKPISVSDVTPYNAVDSSWSSYNEATSAAHYVERDRFVTDL 975
XX	DR; N-PSDB; AA159511.	Db	664 YHTTDPVQATASQGATKPISVSDVTPYNAVDSSWSSYNEATSAAHYVERDRFVTDL 723
XX	Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.	Qy	976 SGHRKHSSTGDLSSFLQQAKNSNLQOREANALAMQKQNSLDGRRHLTUNLISKEIEL 1035
XX	Example 6; SEQ ID NO 3500; 10078pp; English.	Db	724 SGHRKHSSTGDLSSFLQQAKNSNLQOREANALAMQKQNSLDGRRHLTUNLISKEIEL 783
PS	The invention relates to human nucleic acids (AAI5798-AAI61369) and the encoded polypeptides (AAM58642-AAM4213) with nootropic, immunosuppressant and cytotoxic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification.	Qy	1036 RNGELQ 1041
XX	Sequence 812 AA;	Db	784 RNGEVK 789
XX	Best Local Similarity 65.4%; Score 4066.5; DB 4; Length 812;	RESULT 10	ADC37580 standard; protein; 1125 AA.
XX	Matches 787; Conservative 2; Mismatches 0; Indels 57; Gaps 1;	ID	ADC37580
XX	Human nucleic acid associated protein, NAAP-47.	XX	ADC37580; (first entry)
XX	Human; nucleic acid associated protein; NAAP; cyrostatic; KW antarteriosclerotic; anticonvulsant; nootropic; neuroprotective; KW cerebroprotective; anti-HIV; anti-allergic; antiinflammatory; KW thymomimetic; gene therapy; cell proliferative disorder; cancer; KW atherosclerosis; neurologic disorder; epilepsy; Huntington's disease; stroke; immune disorder; inflammatory disorder; AIDS; allergy; KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection.	AC	AC
XX	Hom sapiens.	DT	18-DEC-2003
XX	Human nucleic acid associated protein, NAAP-47.	XX	XX
XX	Human; nucleic acid associated protein; NAAP; cyrostatic; KW antarteriosclerotic; anticonvulsant; nootropic; neuroprotective; KW cerebroprotective; anti-HIV; anti-allergic; antiinflammatory; KW thymomimetic; gene therapy; cell proliferative disorder; cancer; KW atherosclerosis; neurologic disorder; epilepsy; Huntington's disease; stroke; immune disorder; inflammatory disorder; AIDS; allergy; KW developmental disorder; Hypothyroidism; Cushing's syndrome; infection.	DB	DB
XX	Hom sapiens.	PN	WO2003046151-A2.
XX	WO2003046151-A2.	PD	05-JUN-2003.
XX	PD 05-JUN-2003.	XX	XX
XX	PD 26-NOV-2002; 2002WO-US038445.	PF	26-NOV-2002; 2002WO-US038445.
Qy	196 MQEALKVLLAEDGSALSRKVVLTVQVQLRPRPQQASKTSIGHVQVLYRASCFKVT 255	XX	XX
Db	197 MQEALKVLLAEDGSALSRKVVLTVQVQLRPRPQQASKTSIGHVQVLYRASCFKVT 60	PR	27-NOV-2001; 2001US-033325P.
Qy	256 KRDEDSSIMQLKEEFSYEALEHRHDQIVHAMEACRISBEQWSLILYGLDAHKSHMQ 315	PR	07-DEC-2001; 2001US-0340477P.
Db	61 KRDEDSSIMQLKEEFSYEALEHRHDQIVHAMEACRISBEQWSLILYGLDAHKSHMQ 120	PR	14-DEC-2001; 2001US-0340562P.
Qy	316 SIIIDLQSQESFAKSVOELTIVLQRTGDPANTNRPHLELIANDNDPAVSPTEQLE 375	PR	18-DEC-2001; 2001US-0342002P.
Db	121 SIIIDLQSQESFAKSVOELTIVLQRTGDPANTNRPHLELIANDNDPAVSPTEQLE 180	XX	XX
Qy	376 NAMAVKVKVHVGLVDFICNYSRKGHETPQPQNPNSKTKSMCRDLRQGGCPRGTNCTFAH 435	PA	(INCY-) INCYTE GENOMICS INC.
Db	181 NAMAVKVKVHVGLVDFICNYSRKGHETPQPQNPNSKTKSMCRDLRQGGCPRGTNCTFAH 240	PA	Baughn MR, Becha SD, Bhatia U, Burford N, Burrill JD, Blake JJ, Chang H, Chawla NK, Elliott VS, Emerying BM, Forsythe IJ, Gandhi AR, Gietzen KG, Gorvad AE, Griffin JA, Hafalia AJA, Ho A, Ison CH, Jackson AA, Jiang X, Jin P, Kable AE, Khare R, LaI PG, Lee EA, Lee S, Lee SY, Li JK, Lu DAM, Pamkumar J, Richardson RW, Sprague WN, Swarnakar A, Tang YT, Warren BA, Xu Y, Yao MG, Yue H, Zheng W;
Qy	436 SQEELEKVRLNKKCINATVTPFLNKVGVNTVTAGNTVISSTETGKIVSTNG 495	XX	XX
Db	241 SQEELEKVRLNKKCINATVTPFLNKVGVNTVTAGNTVISSTETGKIVSTNG 300	DR	WPI; 2003-513642/48.
Qy	496 ISNAENSQSLRSRSTDSTLRALETVKVKVKGANGONAAGPSADSVTENIGSPPKTPV 555	DR	N-PSDB; ADC37640.
Db	301 ISNAENSQSLRSRSTDSTLRALETVKVKVKGANGONAAGPSADSVTENIGSPPKTPV 360	XX	XX
Qy	556 SVAATSAGPSNVGTELNSVQPKSSPFLTRPVYPHSENTOYFODPRTQFPEVPOYQO 615	XX	XX
Db	361 SVAATSAGPSNVGTELNSVQPKSSPFLTRPVYPHSENTOYFQDRTQTPEVQYQO 420	XX	XX
Qy	616 TGYYPPPTVPGAYAPCPYRFYRSNNVYESSUPPASMYADHYSTSFRDRMNSSPYQPP 675	XX	XX
Db	421 TGYYPPPTVPGAYAPCPYRFYRSNNVYESSUPPASMYADHYSTSFRDRMNSSPYQPP 480	XX	XX

PR New human nucleic acid associated proteins (NAAP), useful for diagnosing, PR treating and preventing diseases or conditions associated with the PR aberrant NAAP expression e.g. cancer, AIDS, atherosclerosis, epilepsy, or PR infections.

XX Claim 1; SEQ ID NO 47; 383pp; English.

XX The present invention relates to novel human nucleic acid associated CC protein (NAAP) (I: Adc3753-Adc3759) and their coding sequences (). The CC NAAPs and their coding sequences are useful in diagnosing, treating and CC preventing diseases or conditions associated with the decreased CC expression or over expression of NAAP, such as cell proliferative (e.g. CC cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's disease, stroke, immune/inflammatory (e.g. AIDS, allergies) and CC developmental (e.g. 9. Hypothyroidism, Cushing's syndrome) disorders, or CC infections. These are also useful in assessing the effects of exogenous CC compounds on the expression of nucleic acid and amino acid sequences of CC XX SQ Sequence 1125 AA:

Query Match 42.9%; Score 2665.5; DB 7; Length 1125;  
Best Local Similarity 51.9%; Pred. No. 9.1e-19;  
Matches 609; Conservative 141; Mismatches 295; Indels 129; Gaps 38;

Qy 1 MPVQAAQTEFLSPCICNEFDENVKHEKISLGSHTCKTCINKLHKACCPDQTANTD 60  
Db 1 MPVQAPQWTDFLSPCICQTFDETRIKRKEISLGCGHTYCKMCLNKLRHKACCPDQTNTD 60

Qy 61 IDVLPVNPAOLQVGAQVDPDHOSTKL-SNLGENKHYEVAKKCYEDLAIYLKPDSGGKVA 119  
Db 61 IELLPVNPSALLQLVGAQVPEQQPTRLCSVEDTPKHYEAKKCVBELLAYLKPKJSSARGV 120

Qy 120 -SINOSLRSRMRQKLUTVLNCOLVEBEGRYRAMRARSGLERTTYTELQHNPQLS 177  
Db 121 LNNTTQSLSRPQRKLUTVLHQLVQLEBEGRITAMRARSGLERTTYTELQHNPQLS 180

Qy 178 :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
Db 181 SNIWAAVRAARGCOPFLGPAMQEALKVIALEDGSALSRKVLVLFVQRLBPFPOASKT 237

Qy 238 SIGHVQQLYRASCFKVTKRDEDSLMLQKEFRSYEARREHDQAQTVHIAEAGLRISP 297  
Db 241 SIGHVQQLYRASCFKVTKRDEDSLMLQKEFRPYEARREHDQSIVQIAEAGLRAP 300

Qy 298 EONSSLLIGDLANKSHMQSIIDKLQSPSPRKSVOELTIVLQRTGDPANLNLRPHILL 357  
Db 301 DQNSSLIGDQSKEKSHMQSIIDKLQTPASFAQSVQELTALQRTGDPANLNLRPHILL 360

Qy 358 ANIDPNPDAVSPTWBOLNAMAVKTVHGLVQIYNQSKRGHETPQOQNNSKYKTSMCR 417  
Db 361 ANIDPSAAPPTEQQLNGLVAVRTVWGLVDYIQNSNSKKGADQDQQFPQHSKKTMCR 420

Qy 418 DLROQGCGRGTNTCAHSQEELEYKRLRNKKINATVTFPLNKVGNNTTAGNVI 477  
Db 421 DMKORGCCRGAQCFKASQEELEKFRKRNKRL--VPRPLASLGQNEGLPSAAIL 477

Qy 478 SVIGSTETGKIVPS-TNGISNAENSVSOLISSTSDSTIRALETVKKGKVGANQNAAG 536  
Db 478 PDEGAVDLSRKPKDPALPNCIVSCTGNTVTQLIPIRPTDP--SYDSSLKRGKI----- 525

Qy 537 PSADSVTENRKIGSPPKTPVSNVAATSAGSNSVNGTELNQVPLKSPFLTRPVV----- 588  
Db 526 --DHLSSAAPGSPDLESVPKSIALPVNP---HSPPRGPADLPMPYTKPLQMV 578

Qy 589 -----YPPHSENTOYFQDFPR-TQIPFEVYQYPTGYYPPPTVAGVPCVPRFVRSNN 642  
Db 579 RGSQLYPAQOTDV-YQDPRGAAPFEPAPYQGSMYTPPP-----QCVSFRV----- 626

Qy 643 PESSLPPASMPYADHYSSTSFRDMNSSPYQPPPQGPVPPVPSGMNAYAPIYDSRRRWR 702  
Db 627 PPPSAPEPAPPYLDHYPL-QERVVNSOYGTQFQQ----YPTI----YPSHYDGRRYRP 677

Qy 703 PPMYQDDIIRSNSLPPMDVNHSSV--YQTSURERYNSLGDGYSTVACQP---PS---E 752  
Db 678 APTSYREBIFESPP-PIEIPPAVPSVYPERCQIEESTYVPAHPTQIRPSYIREP 736

Qy 753 PRTTVPLPRPBCGHKLTSCEBOIRKPDQWQYHTOKAPLYNSSTLPVATQSPTPPSPLS 812  
Db 737 PYSLRPPPPQ----NPSLDLHRKRKEIMOLEERK--VISPPPPA-PSPTLP-PTFH 787

Qy 813 VDFRADFSESVSGTKFEBDHLSHYPSPMCGTIGSCINAIIDSSPQDVITANSNALMDLDSG 872  
Db 788 PBFBLDDLKYAG-KYKGNDYQSYSPWCDTIGSY/GTKDAKPDVTAAGSVEMMNVESK 846

Qy 873 DVKRVHLFETORRKE--EDPILIPFDGPIKTMGAISRSRTGYHTTDVQATASQGS 930  
Db 847 GN--RDRDLDLQRRAAFTSDDLPPEPDRPTVSFGAIRSRSKTIYQAGPMQMAQGA 904

Qy 931 ATKPIVSYSDYPPVNAVDSRM - SYG- -NEATSSAHYVERDRFIYTDLSGHRK- -HSST 984  
Db 905 PTKSINTSDSY- -GTHGGMCASTPSPHONTASQGHFSERERISMSSEVASHGKPLPSAE 962

Qy 985 GDLSLSLQQ---AKNSNLL-LQREANALAMQ-----KNSNLDEGRHLTNL-- 1028

Qy :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:  
Db 963 REQLRLELQQLNHOISQOTQLRGLEANTEALQSQPPPPKPGMISSQQQLSLHLQ 1022

Qy 1029 LSKEIELRNLNGELOSDYTEDATDKPDRDIELSALDTE -- PDGQSEPIERBLDIOLGI 1086  
Db 1023 VBERIGKTRTRELSMB-NQCSLDMK-----SKUNTSKOALENGQBPQNKPABDTL 1072

Qy 1087 SSQNDQLINGMAYENGHPVQOHKEPPKQKQKSL 1120  
Db 1073 TFSSD-----VPGNSALTQENISLISNKTSSL 1099

RESULT 11  
ID ADFF55448 standard; protein; 1109 AA.  
XX ADFF55448;  
XX DT 12-FEB-2004 (first entry)  
XX DE Human novel polypeptide #12.  
XX KW human; brain disease; mental disorder.  
XX OS Homo sapiens.  
XX PN JB2003245081-A.  
XX PT 02-SEP-2003.  
XX PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.  
XX DR WPI: 2003-8571164/80.  
XX PR 25-FEB-2002; 2002JP-00047501.  
XX PT 25-FEB-2002; 2002JP-00047501.  
XX PA for treating mental disorders.  
XX PS Claim 4; SEQ ID NO 12; 401pp; Japanese.  
XX N-PSDB; ADFF55399.

CC The invention relates to a DNA which encodes a novel polypeptide. A CC vector containing the DNA is useful as a reagent in estimation of CC standard substance. The antibody is useful for detecting the vector CC containing the DNA and for screening substances and compounds that CC interact specifically with the vector containing the DNA. The vector CC containing the DNA is used as a pharmaceutical, in treatment or as CC preventive agent with respect to disease e.g., brain diseases preferably CC mental disorders. The present sequence represents the amino acid sequence



allergies, neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocrine disorders, and infections. The proteins can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties.

Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein fragment referred to in the disclosure of the invention

Sequence 421 AA;

Query Match 33.9%; Score 2106; DB 4; Length 421;  
Best Local Similarity 99.2%; Pred. No. 1.7e-15;  
Matches 395; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 644 ESSLPPASMPYADHYSTSFRDRMNNSPQQPPQQPYGVPPVPSGMVAPVYDSSRTRWRP 703  
Db 1 ESSLPPGSMMPYADHYSTSFRDRMNNSSQQPPQQPYGVPPVPSGMVAPVYDSSRTRWRP 60

Qy 704 PMYQDDIIRTSNLSLPPMVMHSSVYQTSEURRTNSLDQSYTACQCPPEPRTTVPLPREP 763  
Db 61 PMYQDDIIRTSNLSLPPMVMHSSVYQTSEURRTNSLDQSYTACQCPPEPRTTVPLREP 120

Qy 764 CGHLKTSCEQIRRKPDQWAQETITQAKLVYSSLPATQSPPTPSPSPLSVDFRADFSESV 823  
Db 121 CGHLKTSCEQIRRKPDQWAQETITQAKLVYSSLPATQSPPTPSPSPLSVDFRADFSESV 180

Qy 824 SGTKFEEDEHLHSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDSGDVKRVLHFET 883  
Db 181 SGTKFEEDEHLHSHYSPWSCGTIGSCINAIDSEPKDVIANSNAVLMDSGDVKRVLHFET 240

Qy 884 QRTKEEDEPIIPEDGPISKWAISRSRTGHTTDVQATASQGSATKPSVSDVVPY 943  
Db 241 QRTKEEDEPIIPEDGPISKWAISRSRTGHTTDVQATASQGSATKPSVSDVVPY 300

Qy 944 VNAVDSSRNSYYGNBATSSAHYVERDRFTVTDLSCHRKSSTGDLSSLEQAKNSNLQ 1003  
Db 301 VNADDSRNSYYGNATSSAHYVERDRFTVTDLSCHRKSSTGDLSSLEQAKNSNLQ 360

Qy 1004 REANALAMQOKWNSLDEGRHLTLNLSEIELRNGELQ 1041  
Db 361 REANALAMQOKWNSLDEGRHLTLNLSEIELRNGEVK 398

RESULT 1.3  
AAM42140 ID AAM42140 standard; protein; 306 AA.  
AC AAM42140;  
XX 22-OCT-2001 (First entry)  
DS Human polypeptide SEQ ID NO 7071.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amytrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

XX Homo sapiens.  
XX WO200153312-A1.  
XX





This Page Blank (verso)

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: August 4, 2005, 01:26:33 ; Search time 26 Seconds  
(without alignments)  
4407.468 Million cell updates/sec

Title: US-10-619-992-2  
Perfect score: 6215  
Sequence: 1 MPVQAAQWTEFLSPICYNE.....MSEDKNDFLPKVANGKVNNS 1191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters:

283416

Minimum DB Seq length: 0

Maximum DB Seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:  
1: Pir1:  
2: Pir2:  
3: Pir3:  
4: Pir4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	843	13.6	1048	T23764	hypothetical prote
2	219.5	3.5	1186	T19050	hypothetical prote
3	207.5	3.3	1198	T49726	hypothetical prote
4	206	3.3	1611	T38236	hypothetical prote
5	203	3.3	786	T01456	extensin homolog F
6	202	3.3	760	T06291	extensin homolog T
7	194.5	3.1	1279	T18312	hypothetical prote
8	193.5	3.1	744	B86255	hypothetical prote
9	192.5	3.1	944	T28134	hypothetical prote
10	191	3.1	368	C29356	hydroxyproline-ric
11	191	3.1	2160	T20241	hypothetical prote
12	190.5	3.1	760	F86387	probable Pto kinase
13	190.5	3.1	842	S60402	protein kinase CLK
14	188.5	3.0	1819	T32008	hypothetical prote
15	183	2.9	280	T03236	extensin precursor
16	181.5	2.9	530	T48627	hypothetical prote
17	181.5	2.9	2282	T42277	DNA-binding protein
18	181	2.9	1257	T00486	serine/chreonine-9
19	179	2.9	2562	T14266	Xin protein - chick
20	179	2.9	2783	A41948	alpha-fetoprotein
21	178	2.9	1013	T33470	hypothetical prote
22	177.5	2.9	416	JU0465	extensin precursor
23	177.5	2.9	620	S0673	hydroxyproline-ric
24	177.5	2.9	1375	S48375	hypothetical prote
25	177	2.8	1812	T49350	breast/ovarian can
26	175.5	2.8	429	T06296	extensin-like prot
27	174.5	2.8	951	T47617	extensin-like prot
28	173	2.8	1262	T25168	hypothetical prote
29	173	2.8	1794	T38439	hypothetical diver

					nascent polypeptid
					alpha-A-crystallin
					extensin - common
					hypothetical prote
					apoptosis associat
					translation initia
					hypothetical prote
					extensin class 1 P
					hypothetical prote
					zinc finger protei
					extensin homolog T
					hypothetical prote
					ascites sialoglyco
					extensin-like prot

#### ALIGNMENTS

					RESULT 1
					T23764
					hypothetical protein M142.6 - Caenorhabditis elegans
					C;Species: Caenorhabditis elegans
					C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
					C;Accession: T27764; T27112
					R;McMurray, A.
					Submitted to the EMBL Data Library, May 1996
					A;Reference number: Z19795
					A;Accession: T23764
					A;Status: preliminary; translated from GB/EMBL/DDJB
					A;Molecule type: DNA
					A;Residues: 1-1048 <WIL>
					A;Cross-references: UNIPROT:O45962; EMBL:Z73428; PIDN:CAA97810.1; GSPDB:GN00021; CESP:M
					A;Experimental source: clone M142
					R;Steward, C.
					Submitted to the EMBL Data Library, September 1997
					A;Reference number: Z20313
					A;Accession: T27112
					A;Status: preliminary; translated from GB/EMBL/DDJB
					A;Experimental source: clone Y52D3
					C;Genetics:
					A;Gene: CESP:M142.6
					A;Map position: 3
					A;Introns: 21/3; 122/3; 381/3; 579/3; 647/2; 683/1; 783/2; 855/1; 909/1; 993/1
					Query Match 13.6*; Score 843; DB 2; Length 1048;
					Best Local Similarity 30.6*; Pred. No. 9.8e-39; Mismatches 99; Conservative 278; Indels 162; Gaps 23;
					Matches 238; Gapped 238; C
					Query 4 OAAQWTERFLSPCIPCYNEFDENYKPIISLGKCSHTVKTCIKNLKH--KACPFDGTDATNTDI 61
					DB 6 QSGQWQBYLCCSICNRHENE-TPLPVSLICGIVCRICKACRKPENQTCPKPHDDWKTHSP 64
					Qy 4 OAAQWTERFLSPCIPCYNEFDENYKPIISLGKCSHTVKTCIKNLKH--KACPFDGTDATNTDI 61
					Db 6 QSGQWQBYLCCSICNRHENE-TPLPVSLICGIVCRICKACRKPENQTCPKPHDDWKTHSP 64
					Qy 62 DVLPVNFALLQLQVGAQVDPDHQSILNLIGENKHYEVAKKCVEDLALYKPLSGKGKGVASL 121
					Db 65 SEPNNTNLLSVT---PFRKQCMNTLSGAVSEAERKVDSLQSI-IAKFPREADSBRG-GTV 119
					Qy 122 NQSALESPPMRQKLVLVNCQLVEBGRVTRAMARSLGERTVTELIQHQNPQQLSANLW 181
					Db 120 SSREISSLQRVTLALLYQWREVDGLKTLKMRGTSERVNIEIILSIIQSNTHVSSQLW 179
					Qy 182 AAVRARGCQFLGPAMQEEALKLVLALAEQDGSALSKRYLFLVQRLPRFPQNSKTSIGH 241
					Db 180 SATVRARGCQFLGPAMQDDVRLRLLMTBGTCAKRNLYMVYVQTLASDYPQSKTCVGH 239
					Qy 242 VVOLYRASCFKTYTKRDESSLMLQKEFERSYEAFLRREHDAQIVHIAEAGLIRISPEQWS 301
					Db 240 VVOLYRASCFKTYTKRDESSLMLQKEFERSYEAFLRREHDAQIVHIAEAGLIRISPEQWS 301

Qy	302	SILYGDILAHKSHNOSTIDKLOSPESFAKSVOELTIVLQRTGDPAN-----INRLRP 352	Qy	439	-----ELEKYRL--RNKKINATVRTEPLINKVGYNNTTAGNVISVGSTETTGKI 489
Db	300	ALLYADQHRSAMQSIIKLQLQSKNSIOQVEERAL--AGSQTSMLVPAFVFLQTQVIP 356	Db	287	STCRSASKLPKDQIPRSPSYTHSLXKRIPITDEPVQKSRSITPPMSHRVRIQSPP 346
Qy	353	HLELLANDPNDPAVSPTWEQLENAMYAVKTVHGLVDFIQNY--SRKGHETPQOPNS 409	Qy	490	VPSTING-----ISNAENSYSQSQLSRSTDSTLRALEYV-----KKVKGKVANGONAA 535
Db	357	CJEFAGAEHDTSMRMIGDALHQIRTLK--LHCSCDDLRKMPKEBERGVTLQAEGVGG 414	Db	347	PESSNSTSYVIPVKDRRSDSPRTV1PFESSNSTTLDKVOSTPTYRISK-----397
Qy	410	KYKTSMCRDLRQGGCPPGTNTCPHSQEELEYRKNNKINATVRTEPLINKVGVNNTV 469	Qy	536	GPSADSUTENKGSPP-----KTPVSNVATAAGPSNVGTLEINSVPQ----KSSP 581
Db	415	-----MGGGFGSG-----GABAGRI-----430	Db	398	-PSQTKEPTNSRVSSPIQSQSYHSMMDKMTSVYRTR-----TNNLKOEPVIVPRTRYSKTP 455
Qy	470	TITAGNYISVGISTEGTKVPSNSTGNGSNAENSVSOLISRSSTDSTLRALEYVKKVGKA 529	Qy	582	FLTRPVYPPHSENIOY--FODPRTOIPFEVQYPTGYYBPPPT-----V 625
Db	431	-----GGIPLPLYSOIDECTERSISPTNPKDNSHSPQ-----461	Db	456	LOKEIPPVPLERSVIROPEYSDPGDRIIVIPFSNOKKYSPPPSNSSTPYHTNLQRDV 515
Qy	530	NGQNAAGGSAADSUTENKGSPPKTPVSNVAATSAGPSNVGTLEINSVQPKSSPFLTRVPPVY 589	Qy	626	PAGVAPCVPFRYRNVPESULLPSPASMPYADHYSTSFSPRDMMNSPQPPPPQQGPVPPV 685
Db	462	-----TPPKQPROKRYQMGIPPNRMG-----YSSDAPPFIPSHQQQ 497	Db	516	PTSPPLPRPLRS-----PTSTASSVE-YIPRAFVPRPLYDREPSTSPRS 565
Qy	590	PPHESENIOYFQDPRTQIPEVPOYPOYQTYPPPTVPGAVACPVPREVRSNNVPESSLPP 649	Qy	686	VPGMY-----APVYDSRRIRRPMPYQDDIRNSLPPDMHSSVYQTSILRER 735
Db	498	PPP---QFFNS---QHLEFQRFGGRQGAPPPQPPQ-----MPMLIGYD-----539	Db	566	BPSEYLTLPLEQEYQVQAPL-SRR----PLYSRPPSTASA PRSP----QPKIYLTQLKE 615
Qy	650	ASMPYADHYSTFSPRDRMNNSP-----YQPPPQPYGVPPVPSGMAYPVYDSRRIRWPP 704	Qy	736	YNSLDGYSSVA-----CQPPSEPRTTVPLR----BPCGH--LKTSCEQERRPKPQWAQ 784
Db	540	APMQATEVLT-ADGQMONGTIPARVVIMQSPTHLPGFPVVMIPQQQMVMPPOSMTVPGP 598	Db	616	VPHVPLYSCVARPRPYSTPPEP--IVIRAVNNPPVSSTYLITTLEPVFGTPLQHYI 672
Qy	705	MYQDDIIRNSNSLPPMDWHSSY-YQTSLRERYNSLDGYYSV--ACQPPSEPRRTVY 757	Qy	785	YHTQKAPLVS-----STLPVATQSPTPPSPLFSDYFDRADSE-SVSGTKFE-----829
Db	599	M-----GPMGMPMTSPIPVQVBPNTMWTATSPTGSVIYTPASPPGQGPHTI 643	Db	673	YRSRSSSTSYVPPVIEASTRTVASPRTYTPPSSSPKQHHLTHLEREVPGTPLNHYSRPR 732
Qy			Qy	830	-----EDH--LSHYSPWSCGTGSCINAIDSEPK-----856
Db			Db	733	STPIYSTPPEDRNVSPSRYTPPSVTMAAKLHLTPQPEVSTPLQHYSRSRASSVY 792
Qy			Qy	857	-DVIANSNAVLMDLQSGDVKRRVHFETQRRTKEEDPPIPSDGPLISKNGAISSSRT 914
Db			Db	793	TPSVDSTSTRSLVSPSSVRMPKLHL---TRLQDVPWTIFK-----VSPRST 838
Qy			Qy	915	GYHTDPUQATASQGSAT-----KPIVSVDVPPVYNAVDSRWSSYNEATSSAHY 964
Db			Db	839	PIYSTPPLKEPLSRTLVTPSRRTPPQKLHQISLQDQIP-----GPLQHY 883
Qy			Qy	965	VERDRFVTT--DLSGHRKHSSSTGDLISLE-----991
Db			Db	884	IYRSRSTSTYSIPLSDYQPQSPPTSSVIMNDSPYBEKPKSYLTGPKTYTRIMPERPPK 943
Qy			Qy	992	-----LOQAKNSLILQOREANALAMOKWNSLDEGRHLTLNLSSKEI--ELRNGE 1039
Db			Db	944	IISPPEHPVPTSKSKSTQTLQDM--VVLPQPKTHLYPEISYRRLPPLSIHVPAEKSST 1000
A; Residues: <WIL>			Qy	1040	LOSDTEDATDTKPDRIEELSALDTDEPDQSEPIE-EILDIQLGISSONDQLINGMA 1098
A; Cross-references: UNIPROT:Q17786; EMBL:Z49908; PIDN:CAA90094.1; GSPDB:GN00020; CESP:CO			Db	1001	IISIYSPIDDVPPTHSELHAKRMIVGIDDELQPEKNVNTVYATAIDNELLHRHP 1060
A; Experimental source: Clone C07E3			Qy	1099	VENGHPV-----QHQKEPPRQKKQSUGEDHVILEEKFTILPVFTSC 1139
C; Genes: C07E3.3			Db	1061	EYGRHPVYVLPQATSYGASAKRSSD-----DEAKNVLPPQNC 1098
A; Gene: C07E3.3			Qy		RESULT 3
A; Map Position: 2			T49726		hypothetical protein B23L21.390 [imported] - Neurospora crassa
A; Introns: 36/3; 79/3; 1128/3; 1160/3			C; Species: Neurospora crassa		
Qy	265	QLKEEFREYEAIR-REHAQVHIAEMGLRESQEWNSLLYGDLAHSRHMQ--SLID--	C; Date: 02-Jun-2000	#sequence_revision 02-Jun-2000	@text_change 09-Jul-2004
Db	59	QLRKREVELETIRQHESREPTIDAYQNQPAQ--GNYLNEFNHYDSMKHASPDA 115	C; Accession: T4976		
Qy	320	-----KLGSPESFAKSVOELTI----VLQRTGDPANIN-----348	R; Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Holland, R.; Nyakatura		
Db	116	TYFALKPPTIEVNQGRQFFSPTTSSTRYDQVLPDPSPEIQRREVPSKOSPPNYPKRSVFG 175			
Qy	349	-KURPHLELLANDPNP----DAVSPTEQLENAMYAVKTVHGLVDFQIYNSRKGHET 402			
Db	176	GRVSP-LPQSEYDPMHMPYRIPPMRPTPKAQYCKTTRV-----TTYTTCG-DT 226			
Qy	403	PQOPNPKY-KTSMCRD--LROGGCPFGT-----NCTFAHSQE-----438			
Db	227	PESPRSTYSSKASKLKEITSVKPYATYSRGTKLUPKYDEVPRNTMYSRASEASAKTSSRTSLA 286			

C:Species:	Schizosaccharomyces pombe
C:Date:	03-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession:	T38236
R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.	
A:Cross-references:	UNIPROT:Q42854; EMBL:AL021813; PIDN:CAA16991.1; GSPDB:GN
A:Experimental source:	strain 97h-; cosmid c23A1
C:Genetic:	
A:Gene:	SPDB:SPAC23A1.17
A:Map position:	1
Query Match Score 3.3%; Score 206; DB 2; Length 1611;	
Best Local Similarity 13.8%; Pred. No. 0.0318;	
Matches 165; Conservative 104; Mismatches 293; Index 272; Gaps 0	
Qy 271 RSYEALRRRHDQIYVHIAEA-----GLRISP-BOWSSLLYGYDL 3	
Db 698 KSFSPRIERKLPSRSEVTDSEIDDKQNEYDPTTSARALPPPLFGLFKVDTLASLAHDL 7	
Qy 309 AHKSHMQSTIDKLGOSPEAKSVOELTIVLQRGT-----PANLNRRLPHI-----BLL 3	
Db 758 DD----LPAPVPRISPPPLPKTPS-----GEFGDNEMFPKKSNNVRGHQSRPSTSQQL 6	
Qy 358 ANIDPNPDAVSPTEQLENAMEVAKTTV-HGL-----VDFIQNYSRKGKHETPQ 4	
Db 808 RNTVPVSTVTSGGRBALPDEMAASPSSIGHPLSPPPADENSLNDYFYEPSYL--ESPA 8	
Qy 405 PQPNSKYKTSMCRDLRQGGCPRTGTCFAHSQEELEYKYLRLNKINNATVTPFLANKVG 4	
Db 866 PEQPSYE-----ESENATVTHA P-----TPSTATFQGHPTISVNA 9	
Qy 465 --VNNTTIVTAGNIVSIVGISTERTGKIVPSTGNSAENNSVOSLRSRSTDTRLATEV 5	
Db 903 TPRPLQDVTSVPSVADASATH-----OSTGTLI-----QBITQLSNN-----MRNP 5	
Qy 522 KKVKGVANGQNAAGP--SADSVTENKIGSPKPKTPVSNVAAAT-----	
Db 946 TKLTRPSNDGRKAQSPRAPPSS-----PPPLPVSNILSSPTEBPKDQHPPSAPLSK 9	
Qy 562 --SAGPSNVGTELANSVPQKSS-----PFLTRVYVYPHSEN 5	
Db 999 PVSTSPAAPIALARPPVKLSSKAPPVPLPSADAPP1PVPSTAPPVPTSTPPVKSSG 1	
Qy 596 IQYFDQPRQTCIPEFYQPYOTGYYPPPTVPGACVCPTRVSNVPPY--LPPAMP 6	
Db 1059 APSAPPVPAVSEPSIPIAPGAPVP--AFSGIPPPVKPSVAAPPVPPKPSVAVPPVPP 1	
Qy 654 YADHYSTFSRDRMNSSPY----QQPPQQPYGPVPPVPSGMAYPYDSSRRIWRPPMYQ " 1	
Db 1118 SG--APPVKPKPSVAPPVPPVPSGAPPVPPKPSVAAPPVPPKPSVAAVVPPV-- 1	
Qy 709 DDI1RSNSLPPMDYMHSSVYQTSLRERYNSLDGYYSVACQPSSE-----	
Db 1172 -PABSSG1PPVKPAAGV-----PPVPPSEAPPVPKPSVGVPPVPP 1	
Qy 753 PRRTVPLPREPGHUKTKSCREQRIRRKPDQWAOQHTOKAPLYS-----STLPV-----ATQOPT 1	
Db 1213 PSTAPVPP-TPSAGLPPVPPVPTAKAPP--VPAPSSEAPSYSTPRSSVSPSPSNASPSPT 1	
Qy 806 PPSPLFSDVFRADPSESVSGTKFEE-----DHLSHSPWNCGTGTGSCINADS 1	
Db 1269 SSMSASAAPARTSVRSKSAERHETSTSSSSKSSRGEBHHHHHN-----	
Qy 854 EPQDVTIANSAVLMDLSDGDUKRRV--HLFPTQRRTKKEEDP1PFSGDPIIISKWGAISRS 1	
Db 1313 --EGHADSSSTRTSLAHQDSRKS1HRLHSRSSSSRASKK-P1IVSTTGPNFNEFSFA-----	
Qy 912 SRTGHTTDVQDATSQ----GSATKPISV--SDIVPVY-----VNADSRNNS 953	

Db	1365	-----KPVEPCASEKWNLNSTAVPKSVYQMDNSVLMIKEGIGTQDKKYKS	1410	
RESULT 5				
	T01456	extensin homolog F24O1.18 - Arabidopsis thaliana		
	C;Species: Arabidopsis thaliana (mouse-ear cress)			
	C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004			
	C;Accession: T01456			
	R;Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Cor			
	C;Cross-references: UNIPROT:Q48809; EMBL:AC003113; NID:92689438; PID:92781362			
	A;Experimental source: submitted to the EMBL Data Library, January 1998			
	C;Genetics:			
	A;Description: Genomic sequence for Arabidopsis thaliana BAC F24O1.			
	A;Number: Z14211			
	A;Accession: T01456			
	A;Status: translated from GB/EMBL/DDBJ			
	A;Molecule type: DNA			
	A;Residues: 1-786 <SH1>			
	A;Cross-references: UNIPROT:Q48809; EMBL:AC003113; NID:92689438; PID:92781362			
	A;Experimental source: cultivar Columbia			
	C;Genetics:			
	A;Map position: F24O1.18			
	A;Note: F24O1.18			
	Query Match Score 3.3%; DB 2; Length 786;			
	Best Local Similarity 23.1%; Pred. No. 0.019;			
	Matches 137; Conservative 58; Mismatches 209; Indels 190; Gaps 30;			
	Db	549 SPPKPKVPSVNAATSAKGPSNVGTTELNSVPQKSSP----PLTRVPPVPHSENIOYFQDPR	603	
	Db	511 SPPPPV-----YSPPPPSPAPTPVCTRPPPPPSPQQFSPPP	554	
	Db	604 TQIPEPVQYQPTGYPPPTVPAQAVCPYPRFVSNNYPPESSLPPASMPYADHYSTFSP	663	
	Db	555 -----PEPYTYSPPPHSSPPHSP-----YGPVPPVPSGMKA----PVYDSRRIRW	594	
	Db	664 RDRMNSSP----YQPPPQP-----YGPVPPVPSGMKA----PVYDSRRIRW	702	
	Db	595 PTPVSSPPPPTPVYSPPPPCIEBPPPPCIEYSPPPVYSSPPPPVYSSPPPPVYSSPPPPV	653	
	Db	703 PPMYQCDDLIRTSNLSLPPDMHSSVYQTSLRERYNSLDGYSYSVACQPPSEP---RTTVPL	759	
	Db	654 PPVY-----SSPPPPPVPYSSPPPEVH-----YHS-----YHS-----PPSPVHSSPPPPP	695	
	Db	760 PREPPGHLLKTSCEBIRRKFQDWQXHTQAKLV--SSTLIPVATQSPTRPSP	809	
	Db	696 PSAP-----CEES---PPPAPVYVHSSPPPPVTHSPSPVTHOSPPSP	737	
	RESULT 7			
	T18312 hypothetical protein L7610.3 - Leishmania major			
	C;Species: Leishmania major			
	C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
	C;Accession: T18312			
	R;Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Rajandream, M.; Iw			
	A;Reference number: Z18876			
	Submitted to the EMBL Data Library, May 1999			
	A;Accession: T18312			
	A;Status: preliminary; translated from GB/EMBL/DDBJ			
	A;Molecule type: DNA			
	A;Residues: 1-1279 <OLI>			
	A;Cross-references: UNIPROT:Q97005; EMBL:AL034256; PID:e13711557;			
	C;Genetics:			
	A;Note: L7610.3			
	Query Match Score 3.1%; DB 2; Length 1279;			
	Best Local Similarity 21.3%; Pred. No. 0.012;			
	Matches 233; Conservative 130; Mismatches 399; Indels 331; Gaps 60			
	Db	80 DHQSILKSLNUGENKHYEV-----AKKCVEDLALYKLPLSGGKGKVASLNOSALS-	127	
	Db	230 DHQF---TFVDHKHHKVQAVDGGALHDANSMSLESSL--MSDGGSFSLNEAIIKLS	284	
	Db	128 --RPMQRKLVLYNCQLVEEGRVTRAAARSLOERTVTELLAQNPOLSLWAAV	184	
	Db	285 MAPRPRRQ---TGPFQFRV---KMRAPHYMOSSQLNRDGFVLSRMTRELVASPECELL	338	
	Db	185 RARGCOF--LGPMAMBEALKLVLL---BDGSLRSRKVLFWQRLEPRPPASKTS	238	
	Db	339 LLRCCEVQIGHNSSELLPAQIAYQDITYLSR--ALLVVRCAIGLCSVYDYD	395	
	RESULT 6			
	exten291 homolog T9E8.80 - Arabidopsis thaliana			

239 IGHVYVOLYRASCFKVTKRDEDDSLMQLKEEFSYEALLREHDAQIVHIAAMEAGURISPE 298  
 Qy | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |  
 Db 396 HFWVTHMLHRA-----DRLLEREEOMRNNNATKQATWK-MHRRAMOO----- 437  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luers, J.S.; Khan, S.; Khaykin, E.; Kim, C.  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Schwarz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 A.; Authors: Salzberg, S.L.; Schwarzbach, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A;Reference number: A86141; PMID:2101619; PMID:1130712  
 A;Accession: B86255  
 A;Species: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-744 <STOP>  
 A;Cross-references: UNIPROT:065375; GB:AE005172; NID:93157926; PIDN: AAC17609.1; GSPDB:GR  
 A;Map position: 1  
 Qy 466 NNTVTITAGNVISVGISSTETTGKIVDSTGNGTNSNAENVSOLISRTSDSTRALETVKRG 525  
 Db 559 RGDTFTFLVNPATVGGTQDFT----QTDTDVHTTEYG-DVLTIVTRTTVTKTIEDL---- 609  
 Qy 526 KVGANQNAAGPSADSYTENKIGSPPPTPVNSVAATSGAPSNGTTELNSVPQKSS---- 580  
 Db 610 ----VDSAGDGTDEBPSEDEGQP----APLHVGA---PPPQRQQQQPR 648  
 Qy 581 -----PELTRVPPYPHSENIOYFQDFPR---TQIPFFEVPOYFQTGYYPPPPTV--- 625  
 Db 649 HATHSTHPOLPEPAP-YPPDDDADETECWSWARAQLQQCQPAPOAPPMPMRE 707  
 Qy 626 -PAGVAPC-----VPPFVRSNNV-----BESLPP-----ASMPY---- 654  
 Db 708 KOTQVHPGHVYTQTTRAPPRVNQTKTAPPRLRREEAEPRTSRQAATTASHPVRET 767  
 Qy 655 ADHYSTSPSDRMNSSPYOPPPOPGP-YPPVPSGRNYAPYDSSRIWRPPMYQR---- 708  
 Db 768 ATHYN--PDRAHHAPAHQPPPAPTPQRQPPVSPS-----PPSCRRAAGS 813  
 Qy 709 ---DDI1RSNLSPPMDMHSSYUQTSLRERYNSLDGYYSVACQPPSEPRRTVPLPREPCG 765  
 Db 814 RPAGGLEADNALXPDTDSTSPTAVAR-ESRENNSLP----ERPPKADPRPVHL---- 861  
 Qy 766 HJKTSCBEEIRKXPDQWA-----QVHTOKAPLVST----LPV---- 799  
 Db 862 HNSRAAENALDEPSSAORGALPQORHTQTAPPAIRS-GAALEPLDDYEDGNAPYQEEA 921  
 Qy 800 -----ATOSPTPPSPPLSVDFRAD----FSESNS-GTKFEDH-----LSH 835  
 Db 922 RSEVIASSDAQPSISSAQSRLYTPDVRENSRRELFDLVEPTNTAFQEBWNGPAFLPLIGH 981  
 Qy 836 YSPWSCGTIGSCINAIDSEPKDVTANSNAVLMLDGSDCDKERVHLFETQRTKEED---- 891  
 Db 982 RRGSGRSQSSOSMPRTSESSGQVOJALHGNGSHPSKQKVPIOLYD-RRHQRGSDCQGQ 1040  
 Qy 892 ---PI---IPSDGPIIISKWMA-----ISRSSRTGYHTDPQO---TAS 928  
 Db 1041 QSGMPMYYLTTPTNEYPLQPAQQPRTVKTRDYKVKTRVRVRYTREPYSIMDSENT 1100  
 Qy 929 GSATKPSIVSDYVPPYNNVADSRWSSYNEATSSAH--YVERDRFIYDLSGHRKHSSTGD 986  
 Db 1101 SSHSGSVSIGE----SMSAHWEDSASVGSSTHVEEDEHDAASVNSKPSN---- 1152  
 RESULT 9  
 T28734  
 hypothetical protein F26G5.9 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T28734  
 R;Submitted to the EMBL Data Library, September 1997  
 A;Description: The sequence of C. elegans cosmid F26G5.  
 A;Reference number: Z20516  
 A;Accession: T28734  
 E86255 hypothetical protein [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
 C;Accession: B86255  
 R;Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar, K.,  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-944 <STOP>  
 A;Cross-references: UNIPROT:O16936; EMBL:AF022974; PIDN: AAC48041.1; GSPDB:GN00023; CESP

A; Experimental source: strain Bristol N2; clone F26G5	Qy 877 RVHLIFETQRRTKEEDPII--PFSDDGPIIISKNGAISRSRRTGYHTDPVQATASQGSATKP 934
C; Genetics:	Db 701 VIHF-----EEVVVLBDVDGRDHANGILENNVEEQQFDD-----VP 740
A; Map position: 5	Qy 935 ISVSQDYYPPYNAVDSRWSSYGNNEATSSAAHYVERDRFLVTDLSGHRKHSSTGDLISLELQQ 994
A; Introns: 12/3; 48/3; 258/3; 327/3; 643/1; 916/2	Db 741 PQINNFVEF-----DDREQVDADEEFP---EPVHLR-----AAIEFQE 777
Query Match 3.1%; Score 192.5; DB 2; Length 944;	Qy 995 AKNSNLLLQQ- REANNLAMOKWNNSLDEG-- RHLTINLSSKEIERNLNGELQSDYTEDATD 1050
Best Local Similarity 17.7%; Pred. No. 0.0096;	Db 778 AADQDALEBODYERNLDAEENEEDGGDLPYEEBDEDEEVHVTYVR-RVQRPRD 836
Matches 215; Conservative 146; Mismatches 410; Indels 445; Gaps 56;	Qy 1051 TKPDRIEELSA-----DTEDPGQSEPIEFLIDOLGISSQNDDLNGMAGENV 1102
Qy 7 QTEFEGLSCPICYNEFEDENVHXPISLGCSHTYCKTCINKLKHAKCPFDQTAINTDIDVLP 66	Db 837 AAHISEVETTRSSARLARNSNPTDEPSTS-----GAAGPSN 874
Db 14 QWAHVVBEPICVNIYD---KEMQMGCGHLCSTCIGRL-----VDQVKT 54	Qy 1103 HPVQQHQKEPPKQKQ 1118
Qy 67 NFALLQLQVGAQPDHOSIKLNLNQGENKHYEAKCVCEDALYKPLSGKGKVASYLNQSA- 125	Db 875 RP--PHSPTPPPGNKR 888
Db 55 NMA-----EEBLGRPR-----	Qy 126 -LSRPMDRKLVTLNCOUVEBEGVRAMAARSILGERTVTLELQHNPQQLSANLWAAV 184
Qy 86 QEDQPHFTDLPFLDPRMNPNDNMNMRGI-----VWLPRRGPE----- 122	Db 887 AHISEVETTRSSARLARNSNPTDEPSTS-----GAAGPSN 874
Qy 185 RARGCQFQGPANOQEBAULKVLLAEDGALSRSKVLFLVYORLEPFRPAQSKTSIGHVQ 244	Db 123 -ARATEIKCPECRKPT----LVPADPLPVNTRVQE--IIVQVKAPIFDR-----HLVK 168
Db 123 -ARATEIKCPECRKPT----LVPADPLPVNTRVQE--IIVQVKAPIFDR-----HLVK 168	Qy 245 LLYRASCFKTVKRDDESSLQMLKEEPBSYEALRREHDQIYHIMAFGLRISPEQWSLL 304
Qy 169 LCNQ---CEAVLSQLGVYPCDSOCBETGRK-----ICSTCAIRU----- 203	Db 169 LCNQ---CEAVLSQLGVYPCDSOCBETGRK-----ICSTCAIRU----- 203
Qy 305 YGDLAHKSHMQSITIDKLQSPESPESAKSVQELTIVLQLTGDPANLRLRPHLELLANTDNP 364	Db 204 ----HRNH-----OLIVEKKALTSDVREMKQKISDASH-----AFAQSLENLKPR 245
Db 204 ----HRNH-----OLIVEKKALTSDVREMKQKISDASH-----AFAQSLENLKPR 245	Qy 365 DAVSPITME--OLENAMYAVKTVVHGLDFDIONYSRKSHETPQPQSKYKXTSMCRDLRQO 422
Qy 246 QSGGTIEAKAQLDQLTLSIKIPBFMULSTF-----DSKIK----- 279	Db 246 QSGGTIEAKAQLDQLTLSIKIPBFMULSTF-----DSKIK----- 279
Qy 423 GGPPRGNCNTCTFAHSQEELEKYLNRKTKINATYRTEFLINKGVVNNTTTAGNVISIGS 482	Db 280 -----ENSTDMEEVTK---AEKSKY----- 301
Db 280 -----ENSTDMEEVTK---AEKSKY----- 301	Qy 483 TETTGKVPSTNGISNAENS-----VSQSLRSRSTDSTLRALEYTKVGRVGANGQNA 535
Qy 302 SPPAGRIVDVMALTAIESFNSYNDPFEEKQEQFD-----EPTGSVNI----- 347	Db 302 SPPAGRIVDVMALTAIESFNSYNDPFEEKQEQFD-----EPTGSVNI----- 347
Qy 536 GPSADSYTNTKGSPPPKTPVPSVNTAATSGAPSNGVTELNSVPQKSSPFLTRVPPPHSEN 595	Db 348 -PASEGGER-----VNEVAAPAG-----RPIPNARNLRGALNHFQAVARN 389
Db 348 -PASEGGER-----VNEVAAPAG-----RPIPNARNLRGALNHFQAVARN 389	Qy 596 IQYFQDRTQIPEVQYQPTGQYPPPTVPAVAPCVRFRYSNNNPESSLPPASMPYA 655
Qy 596 IQYFQDRTQIPEVQYQPTGQYPPPTVPAVAPCVRFRYSNNNPESSLPPASMPYA 655	Db 390 -----RNQARGVVOOQDLPFLPPPPP-----HQLYMPPPAMP 424
Db 390 -----RNQARGVVOOQDLPFLPPPPP-----HQLYMPPPAMP 424	Qy 656 DHYST---FSPR-DRNSNSPYQP-----PPPPQYGPVPPVPSGMYAPV 694
Qy 656 DHYST---FSPR-DRNSNSPYQP-----PPPPQYGPVPPVPSGMYAPV 694	Db 425 QHNOAQRLWGGPQQRQIN-----QMMACQGGGRGQYIMGHPPMHQMMQPRMNY-M 478
Db 425 QHNOAQRLWGGPQQRQIN-----QMMACQGGGRGQYIMGHPPMHQMMQPRMNY-M 478	Qy 695 YDSRRRIWTFPPMYYDLDITRSNSLPPMDMHSSVYQTSLSRERYTQVACOPPSEPR 754
Qy 695 YDSRRRIWTFPPMYYDLDITRSNSLPPMDMHSSVYQTSLSRERYTQVACOPPSEPR 754	Db 479 HQHH---PMHOVED-----QMORQHMRQMPNN-----NIMHGMMQIAFQHMHN 523
Db 479 HQHH---PMHOVED-----QMORQHMRQMPNN-----NIMHGMMQIAFQHMHN 523	Qy 755 TTVPPLPREPCGHKTCSQEQRKPDQWQY-----HTQKAPLVSSTLPV 800
Qy 755 TTVPPLPREPCGHKTCSQEQRKPDQWQY-----HTQKAPLVSSTLPV 800	Db 524 ANYMPMPMPMPAHFQ---QQQQQRHPNEQQFOGPPPPPPQRTQNHQOMOQYHQQLQVQ 580
Db 524 ANYMPMPMPMPAHFQ---QQQQQRHPNEQQFOGPPPPPPQRTQNHQOMOQYHQQLQVQ 580	Qy 811 TQSPPTPP-----SPLFSTVDFRA-----DF-----SESYSGTCFEEDHLSHYS 838
Qy 811 TQSPPTPP-----SPLFSTVDFRA-----DF-----SESYSGTCFEEDHLSHYS 838	Db 581 QQQPQPQPAQPFRRPPIQQQQLQDIDAMNDRQFQQLADEVAEDQMVQADPAPAD 640
Qy 819 -----W-----SCRTIG-----SCINAIDSEPKDV-----TANSVAVLMDLDSGPVKR 876	Qy 788 OKAPLYSSTLPTVATOSPTPPSP 809
Qy 819 -----W-----SCRTIG-----SCINAIDSEPKDV-----TANSVAVLMDLDSGPVKR 876	Db 297 PPPPSPPPPPPYKSPPPPPSP 318
Db 641 IPRADWEIQLQVAFORGAGDFGINFHPEPVEEVEQPPVQVQVNNQCVVENVYRAGDINN 700	



Qy	531 GONAAGPAGDSVYENKIGSPPKTPVSNVAATSAAGPSNNGTELNVRQPKSSPL-----	583		Db	90 SYVALYRNKDQNDDAKIPLTSISIVSRTOLKQYCF-ELVRCSDRNSV-----	137
Db	35 GDRATSPSPREPTN---GNNPPE---TNTPAQSSPPPE-TPISSPPSPSPSLTGPP	86	Qy	329 KSVQVELTIVLORTGPDLANLRPHIELLANIDPNPDASPTWELENANAYAVKTV-VH	86	
Qy	584 -TRVPVYYPHSENIIQYFCDPRTQIPFEPV-----QYQPTGYYPPPTVPGAVAPCPV	634	Db	138 -----SSGSSSSLN-----VSSDSNSKCKSIVIATKCESDLH 168		
Db	87 PTIPIPSSPPPEPS----PPEPLPTAAPPANPVSSPPPESSPPPEBAPPPTPITS 140		Qy	387 GLVDFI-----QYSRKSHETPOPQNSRYKTSMSMCRDLRQQGSCPRGTNCFT 433		
Qy	635 RFVRSNNYPPESLPPASMYADHYSTSPDRMANSSPYQPP---PPQPYGPVPPVPSGMY	691	Db	169 SWLDAFAKCPILGSYSSPNTFHVKHVGDPETGSFV-----GMPTWELL 216		
Db	141 PSPTPNPPPPSPES-SIAPDPPSNPLPPKLVPSPSHSPRHLPSPPASETPPPP---	195	Qy	434 AHSQEELKEYLRLRNKINATVRTFPPLINKGVNNNTAGNVISVGSTETTGKLV	PST 493	
Qy	692 APYYDSKRRIWRPEPMYQDDDIITRSNLSLPPMDYKHSYYQTSLERY-----	736	Db	217 KHS-----RITGEDN-----NNSA-----VIVQL-----QPYQEY 243		
Db	196 -----RLHPSPPASER-----STPSDSEHPSPSPGHPKRREOPPPGSKRPTPSP	243	Qy	494 NGISNAENSQSLISRSSTDSPTRALTV--KKVKGKVANGQNAAGPSADSVTENKIGSP	550	
Qy	737 -NSLDGYWSVACOPPSEPRRTYPLPREFGHLKTSCEBQIRRKPDQWAQYHTQKAPIVSS	795	Db	244 NGAGNPNTNLQKPOSGETSSSQKSPLPSYNNDKLNNSVNNSVSSSVSSMVSORKTQSP	303	
Db	244 PSPSDSKRPKVHPSPPSPDEETLPPPKP-----SPD-----PLPSN 278		Qy	551 PKT-----PVSNAATAATSGPSNVSTELNSVPLQKSSPFL-----TRVP 587		
Qy	796 TLPVATQSP-----TPPPPLFSVDRADFSSEVSG 825		Db	304 PNTKSPVSLGSSCSSLPPINTLPLTS--QSNIRHLQVNPQYKRNNGHSTNGQPRGP	361	
Db	279 :SSSPPTLLPSSVSVSPSSPP-----RKSVSG 304		Qy	588 VYPPHSE-NIQYFQDPRTQIPFEPVQYQP-----QTYGYPDPPTVPGAVACPVRVSNVPE	644	
Db			Db	362 MHPNNSORSLSLOCQQQQQQQQQQQQQHQHQQPYHHQGPSPSPSPSPSPNPRPHNMN--PY	419	
			Qy	645 SSLPPASMPYADHYSTSPDRMNSSP-----YQPQPPQPYGPVPPYPSGMYAPVYDSRR	700	
			Db	420 SKQPQSPLSQSTQNQAIPIRQAQNSSTAAHFQ-----PQRTAPKPKPS-----AP	465	
			Qy	701 WRPPHYORDDIIRSNSLPPDMVHSSYQTSURERYNSLDGYVSACQPPSEPRRTVPLP	760	
			Db	466 -RAPYPPBNQNTSNTNHQVAPKNDQSTPQTMR-----QAPKRPDADYAQP	510	
			Qy	761 REPGCHLKTSCBEQIRRKPDQWAQYHTQKAPLVSSTLPVATOSLPPFSVDFRADFS	820	
			Db	511 ----GGV-----AKPKKPARPTMSTAEMSUKKXVTNA--DPSQCFCFKVIBKA-G	553	
			Qy	821 ESVGTKF-EEDHLSHYSPNSCGTIGSCIMADSP-----KDVIA 860		
			Db	554 QGASGSVYLAERTHI----PTESNMELINNIDD-EPHVGDKRVAKOMVLRQPRRELIV	608	
			Qy	861 NSNAVLMDDSDDVKRVVHLFPTQRTRKEE-DPIIPFSDGPILISKXGAISRSSRTGYHT	919	
			Db	609 NEILVMKD--SRHRQIVNLEAYLRTDDLWVWMEMEG----GSL-----T	649	
			Qy	920 DPVQATASQGATSKPISVSDVPPVNAVDSRWSYNEATSSAHYVERD	968	
			Db	650 DIIENSPTINDNSHSPLT-EPIQIAYI---VRETQGLKFLHDKHIIHRD	693	
				RESULT 14		
				T32008		
				hypothetical protein K10G6.3 - Caenorhabditis elegans		
				C;Species: Caenorhabditis elegans		
				C;Accession: T32008		
				R;Holdmann, P.; Mullen, G.		
				R;Davidson, S.		
				Submitted to the EMBL Data Library, July 1997		
				A;Description: The sequence of C. elegans cosmid K10G6.		
				A;Reference number: Z21111		
				A;Accession: T32008		
				A;Status: preliminary; translated from GB/EMBL/DDBJ		
				A;Molecule type: DNA		
				A;Cross-references: UNIPROT:P48562; EMBL:U23084; NID:91050853; PID:9105		
				R;Maurer, K.C.T.; Urbanus, J.H.M.; Planta, R.J.		
				Submitted to the Protein Sequence Database, April 1996		
				A;Reference number: S63266		
				A;Molecule type: DNA		
				A;Cross-references: ENB:Z71574; NID:91302389; PID:CAA96216.1; PID:e239731; PID:9130239		
				R;Cvrckova, F.; Nasmyth, K.		
				Submitted to the EMBL Data Library, November 1994		
				A;Description: S7E20-Like protein kinases are required for cytokinesis.		
				A;Accession: S53103		
				A;Molecule type: DNA		
				A;Residues: 1-842 <MAW>		
				A;Cross-references: ENB:X02499; NID:9732943; PID:CAA57879.1; PID:9732944		
				A;Experimental source: strain K1107		
				C;Genetics:		
				A;Gene: SGD:CL44 ; ERCl0		
				A;Cross-references: SGD:S0005242; MIPS:YN1298W		
				A;Map position: 14L		
				C;Superfamily: plectrin repeat homology: protein kinase homology		
				C;Keywords: ATP; phosphotransferase		
				F,544-825/Domain: protein kinase homology <KIN>		
				Query Match 3.1%: Score 190.5; DB 2; Length 842;		
				Best Local Similarity 19.9%; Pred. No. 0.01; Mismatches 275; Indels 237; Gaps 37;		
				Matches 153; Conservative 104; A; Map Position: 2; A; Intron: 41/2; 191/2; 556/2; 701/2; 1645/2; 1700/2; 1768/1; 1797/3		
				Qy 272 SYEAL--RREHDAQIVHIAAMEAGLRSIPEQNSLILQKQSPEFA 328		



This Page Blank (uspto)

APPLICANT: Hilbun, Erin  
 APPLICANT: Donoho, Gregory  
 APPLICANT: Turner, C. Alexander Jr.  
 TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
 FILE REFERENCE: LEX-017-USA  
 CURRENT APPLICATION NUMBER: US/09/854,856  
 PRIOR APPLICATION NUMBER: US 2001-05-14  
 PRIOR FILING DATE: 2000-05-19  
 NUMBER OF SEQ ID NOS: 64  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO: 48  
 LENGTH: 1939  
 TYPE: PT  
 ORGANISM: Homo sapiens  
 FEATURE: VARIANT  
 NAME/KEY: LOCATION: (1) ..(1939)  
 OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-854-856-48

Query Match 3.1t; Score 193.5; DB 4; Length 1939;  
 Best Local Similarity 19.2t; Pred. No. 7.3e-05;  
 Matches 256; Conservative 168; Mismatches 527; Indels 379; Gaps 57;

Qy 114 GKGVAISLNOSALSRPMQRKLYLVNCLOLVEERGRVR-AMRAARSIGERTTYTELILQHON 172  
 Db 650 GQSSSSSTGVSQSQPIHRPOQQQGIOOTAPPQTQVNSLSQTSSBATTAAQPVSPPQA 719  
 Qy 173 PQQLSANIWAABRGCCFLGPQ-...MQEALKLVIALEDGSALSRSKVLVLFVYQLR 227  
 Db 720 PQVLPQVSAKGQKSTQGVSQVAPKEPVAVAOQPTOPTCLASSYDSSAHSDF--VASGNSDG 776  
 Qy 228 EPRPQOACFTRKDEDSLMKL-...EEFRSVEALRREHD 281  
 Db 777 NEVPSSSGRHEGRTTKHYRK-VRSISRHEKTSRPRLIRLNVSNGDRVTECQLETHN 835  
 Qy 282 AQIVHIAAEAGLRISEPWQSSLL-YGDLAKHSHMQSITDKLOSPESPAKSVOBLTIVLQR 340  
 Db 836 RKNVTFKFDL-DNDNPBEATIMVNNDPILATERESFFDQVREBIEKADEMUSEDNSVSEP 894  
 Qy 341 TGDPAANLRRLPHELLANID-...-PNPDAVSPTEWQLEJAMAVKTY 384  
 Db 895 EGDOG-...-LESUQKDDYFGSGSKOLEGEFKQPIPASMP--QOQIPTSSITQV 944  
 Qy 385 VAGLVDFIONYSRKGHEQPQPN-...-SKYKTSMCRDLRQGCCPRTNCTPAHS 436  
 Db 945 VH-...-SAGRPFIVSPVPESRLLRESKVPSBEITDVAASTAQSPG-...-MNLSHS 991  
 Qy 437 QEELEYKR---LRLNKI-...-NATVRTFP---LLNKVYVNNTVTTPA---473  
 Db 992 ASSLSLQAFSEIERRAQNTEGPNTAPPNFSHTGPTFPVPPFLSSIAVGPTTAATAAPVP 1051  
 Qy 474 ---GIVVISVGSTETGKIVUPSTNGISNAENS-...-VSQLIS 508  
 Db 1052 ATSSPDNLSTVQSE-...-VPTERGIAFGATVSTGVTSGGLPIPVSSEPVLSVS 1108  
 Qy 509 RSTDSTLRALETVKVGKVGAN-...-GQNAAGPSAD-STVENKIG---SPPKTPVSN 557  
 Db 1109 SITIPAVVISISTSPSLQVPTSTSEIVVSTALYPSVTSASAGGSTATPGPKPEAV 1168  
 Qy 558 VAITSAGPSSNVGTELNSV-...-PQKSSPPLTRYVYYPHSENTOYF 599  
 Db 1159 VSOQAAAGSTIVGATLTS-STTTSPPTSTASQLS1QSSSTSTPLAETVVSASHSLD-KTS 1227  
 Qy 600 QDPRTQIIPFVQPYQPYQTYGYPPPP-...-TVPAVGAVPC-PRFVRSNNVPESSLPPAMP 653  
 Db 1228 HSSTTGFLFSL-SAPSSSSSPGAGVSSYTSOPCGLHPVIPSASTSTLPOAGPSTP 1286  
 Qy 654 YADHYSIFSFRDRNINSSSYQPPPPQPYQGPVPPVPSGM-----YAPYDS 697  
 Db 1287 L-...-LPQVFSIPLVQPYQVANPAVQQT1LHSQOPQPLFNQPHTHCPVEVDS 1334

Qy 698 RRIWRPPMYQRDDI----IRS--NSLPPMDVMHSSV-YOTSLRERYNSLDGYYSVA-- 746  
 Db 1335 DT-QPKARGIDDIIKTLEEKLRSLSFEHSSSGAQASVSLSETSLVIESTVTPGPIPTAVA 1392  
 Qy 747 ---COPPSE-PRTTVPLPREPC---GHLKTSCEBQIRR-KPDOWAQHTQKA 790  
 Db 1393 PSKLLTSTTSTLPLTNPLGTVLPAVTPVTPGQVSTPVSTTSGVYKPGT---APSXP 1448  
 Qy 791 PLVSS-TLPVATQSPTPPSPLFSVSDRADFSESVSCTKFEDDHLSHYSPWSGTIGSCIN 849  
 Db 1449 PLTKAPVLPVGTELP-----AGTLPs--E 1470  
 Qy 850 ALDSEPKDVIANSNAVLMDLDS-----GDUVKRVRHLFETORTK--- 888  
 Db 1471 QLPPFPGPSPUTQSQQPLLEDQALQRLRTLSPPEXITVSAVGPVSMALAPTAITEAGTQPKQKG 1530  
 Qy 889 ---EBDP1IPPSDGP1IISKNG---AISRSRSTGTYHTDPVQATASQGATKPSVS 938  
 Db 1531 VSQVKEGPVLATSSGAGVFKRGMRFQVSVAADGAQKGKNSDEAKSVHFESETSSESSVLS 1590  
 Qy 939 DRYPVYNAY-----DSRWSSYGNBEATS-SAHYVZEDRFVTDLS---G 977  
 Db 1591 SSSPESITLVKCPBPNGITIPGTSVSDPVAHKTASAKDPSQPTKVGRFQVTTANKVG 1650  
 Qy 978 HRKHSSTCDL-----LELQOAKNSNLLQREANALAMQKRN-----1016  
 Db 1651 RFSVSKTEDKTDTKTKEGVSPSPFMDEQAVLPAPIPKKCKPPELSEPSHLNGPSSDEPA 1710  
 Qy 1017 ---SLDegr --HILTFLNLISKEIELRN---GFLQSDY-TEDATDTKEDRDIDEL 1061  
 Db 1711 AFLRDVDGGSSPHQLOSSKSLPSQNLSNSFNSNMSDNEDIEDDLKEL 1770  
 Qy 1062 SALDTDEPDGQSEPIBILDTQOLGISSONDQLN-----GMAVEVNGHPVQHQKEPK 1114  
 Db 1771 RRL-----RDKHLKEDIQDLSQRKHIESLYTKLGKVPPAVIIPPAAPLGRRRRPTK 1823  
 Qy 1115 QK---KQSLGEDH-----VILEEQTILPVTSQPLPVJSINASCLP 1155  
 Db 1824 SRGKSSSRSSSLGKNSQLSGQSAASVLPHQQTLLHPPSN-----IPBSGQNONLQP 1878  
 Qy 1156 ITTSVSAGNL 1165  
 Db 1879 LKPSPSSDNL 1888

RESULT 9  
 US-09-854-856-16  
 Sequence 16, Application US/09854856  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walk, D. Wade  
 ; APPLICANT: Hilburn, Erin  
 ; APPLICANT: Donoho, Gregory  
 ; APPLICANT: Turner, C. Alexander Jr.  
 ; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides  
 ; FILE REFERENCE: LEX-0178-USA  
 ; CURRENT APPLICATION NUMBER: US/09/854,856  
 ; CURRENT FILING DATE: 2001-05-14  
 ; PRIORITY APPLICATION NUMBER: US 60/206,015  
 ; PRIORITY FILING DATE: 2000-05-19  
 ; NUMBER OF SEQ ID NOS: 64  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO: 16  
 ; LENGTH: 1999  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1). .(1999)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-854-856-16

Query	DYVPPYNAV-----DSRNSSYGNBATS-SAHYVERDRFLVTDLS--G 977	Qy	939	DYVPPYNAV-----DSRNSSYGNBATS-SAHYVERDRFLVTDLS--G 977
Best Local Similarity	3.1%	Score	193.5;	DB 4; Length 1999;
Pred.	No. 7.7e-06;	Mismatches	168;	Indels 379; Gaps 57;
Matches	256;	Conservative	168;	Db
Qy	114 GGKGVASILNQALSRLSPRMQRKLVLTVNLNCQLVESEGRVB-AMRAARSLGERTVTTELLIQHQN 172	Qy	1651 SSSPESTIVKREPNGITIPGISSDVBSAHTTASAKDTQPTKVGRQVTTANKVG 1710	
Db	720 GQPSSSSLTGVSQSQQPIHQPOQQQGICATPQTVOYSLSQTSTSSEATAAQPVSQPQA 779	Db	978 HRKHSSTGDLSS-----LELOOAKNSNLORANALAMQQWN-----1016	
Qy	173 PQOLSANIIWAAYVARGCCFLGP-----MOEBALKLVLAEDGSALSRKVLYLQHQN 172	Db	1711 RFSVKRBDKLTDRKGPAVASPPFDLEQAVLPATPKKEKPEJEPSPHLNGPSSDPEA 1770	
Db	780 PQVLPQVAGKOSTQGYSQVAPVAAPEVAAQPAQTOPTLASSVSDAHS-----VASGMSDG 836	Qy	1017 ----SLEDGR--HUTINLISKETELBN-----GELQSDY-TEDATDKPDRDIELEL 1061	
Qy	228 EPRFOASKTSIGHVYVOLYRASCFKVTKRDDESSLIMQK-----EEFRSYEALRREHD 281	Db	1771 AFLSRDVDDGSSPHOLSSKSPLSONLSNSNSFNSSYMSNSDEDEDIURKLEL 1830	
Db	837 NEVPSSSGRHEERTTKHYRG-----VRSRSHERTSRPLRLINVSNGDRVVECQLETHN 895	Qy	1062 SALDIDEPDGSESEPIBILDQLGISQNDOLLN-----GMAVENGHPVQHQKEPPK 1114	
Qy	282 AQIVHIAAEGLRISPECWSSLLY-GDLIAHKSHMQSITDKLOSPESFAKSVOELTIVLQR 340	Db	1831 RPL-----PKDKHLEBIQDLSRQKHEIESUYTUGKVPPAVIIPPAAPLSGRERRPTK 1883	
Db	896 RQMVTFKELD-GDNPELIATIMVNNDFILEARESPVDOVREIEKADEMSEDVSVEP 954	Qy	1115 QK-----KOSLGEHD-----VILEEKXKILPVTSFCSQLPLPVYSISASCLP 1155	
Qy	341 TGDPAVSPTEWOLENAWAVKTV 384	Db	1884 SKGSKSSRSSLGKSPQLSGNLGSOSAASVHPPQTLHPPGN-----IPBGGQNQLLQP 1938	
Db	955 EGDOG-----LESLGKDDYGPFGSSQKLEGEGFOPAPIPASSMP-QQIGPTSSLTQV 1004	Qy	1156 ITTSVSGNL 1165	
Qy	385 VHGLVDFTIONYSRKRGHETPOPQEN-----SKYKTSMCRLRQOGCPRTNTCFAHs 436	Db	1939 LKPSPSSDNL 1948	
Db	1005 VH-----SAGRFRIVSPVPSRSLRESKVFSEITDVAASTAQSG-----MNLHS 1051	RESULT 10		
Qy	437 QEELEYKR---LRNKKT-----NATVTRFP-----LINKVGNNTVTTA-- 473	US-09-854-956-46		
Db	1052 ASSLSLQAFSELRLRAQMTEGPNTTAPPNSFHSTGPTFVVPPLSSLTAGVPTAAATAPVP 1111	Sequence 46, Application US/09854856		
Qy	474 -----GRVISVGSTETGK1VPSTINGISNENS-----VSQTLIS 508	Patent No. 6541252		
Db	1112 ATSSPPNDISTSVIQSET-----VPTEBGIAQVATSTGVVTSGGLP1PPVSESPVLSVVs 1168	GENERAL INFORMATION:		
Qy	509 RSTDSTIRALETYKVKGVGAN-----GONAAGPSAD-SVTEENKIG----SPPKTPVSN 557	APPLICANT: Walk, D. Wade		
Db	1169 SITIPAVVSI1STSPSLQVPTSEIVVSSATLYPSVTSATSASAGCTTAPGPKEPAV 1228	APPLICANT: Dorobo, Gregory		
Qy	558 VAATSAGGSNVNGTELNST-----PQKSSPFLTRVPPVYPVPHSENIOYF 599	APPLICANT: Turner, C. Alexander Jr.		
Db	1229 VSQOAAGESTTVGATLTSVSTTSFPTASQSLTQSSSSTPTLAETVVSAAHSLD-KTS 1287	TITLE OF INVENTION: Human Kinases and Polynucleotides		
Qy	600 QDPRTQIPEEVPPVQTYPPPPP-----TVPGAVAPC-----PFRVRSMNPBPSSLPPAMP 653	FILE REFERENCE: LEX-017 USA		
Db	1288 HSSTTGFLASF-SAPSSSSSPGAGVSSYISOPFCGLHPIVIPSIASTPQAAAGPTSTP 1346	CURRENT APPLICATION DATE: 2001-05-14		
Qy	654 YADHSTSFRDRMNSSSPYQPPPPQPYCPVPPVPGSM-----YAPVDS 697	PRIORITY NUMBER: US 60/206, 015		
Db	1347 L-----LPQVSSPILVQPVANPAVQVNTSPTLAEVTSVVAHSLS-D-KTS 1287	PRIORITY FILING DATE: 2000-05-19		
Qy	698 RRIWRPPMNYQRDDI-----IRS-----NSTLPPMDVMHSSY-YQTSLRLERYNSLDGYYSA-- 746	NUMBER OF SEQ ID NOS: 64		
Db	1395 DT--QPKAGIDDIKTLEBEKLRSLSFEHSSGAQHASYSLETSLVTSVTPGIFTAVA 1452	SOFTWARE: FastSEQ for Windows Version 4.0		
Qy	747 -----COPPSE-PRTVPLPREPC-----GHLKTSCEBOIRR-KPDQWAQYHTQKA 790	SEQ ID NO: 46		
Db	1453 PSKLLTSTSTCLUPPTNPLGTVALPVVTPGQVSTPVSTTSGVYKPGT-----APSXP 1508	LENGTH: 2076		
Qy	791 PLVSS-TLPVATQSPTPPSPLFSDVFRADFSESVSGTKFEEDHLSHYSPWSCGTIGSCIN 849	TYPE: PRT		
Db	1509 PLTAKPVTIVGFTELP-----AGTLPs-E 1530	ORGANISM: Homo sapiens		
Qy	850 AIDSEPKDVIANSNAIMLDSL-----GDIVKRRVHLEFQRTK--- 888	FEATURE: Other Information: Xaa = Any Amino Acid		
Db	1531 QLPPFPGPSSLTQSCQPLLEDQAQLRRTISPEXITVTSVAGPVMSMAAPTATTEAGTQDQKG 1590	US-09-854-956-46		
Qy	889 -----EEDP11PFSDGP1SKWG-----AISRSRRTGHTTDBQVATASQGSAATK1SWS 938	OTHER INFORMATION: Score 193.5; DB 4; Length 2076;		
Db	1591 VSQVKEGGPVMLATSGAGVFKMGREFQVSYAAGDQKECKNKSDEAKSYTHESESSESVLS 1650	Qy	114 GKGKVASLNQSLSPRMORKLVLYNQOLVEEGRVR-AMRAARSGLGERTVTELLQHQN 172	Qy
Qy	828 AQIVHIAAEGLRISPECWSSLLY-GDLIAHKSHMQSITDKLOSPESFAKSVOELTIVLQR 340	Db	660 GQPSSSLTGSSQSOPIQHPPQQQGIOQTAPOQTOQYSLQTSQTSSSEATTAQPVSQPQA 719	
Db	836 RQVTFKFPLD-GDNPREIATMVNNDFILATERESPVDFQVREIIEKADEMSEDVSEPV 894	Qy	173 POOLSANIWAARVARGCQFLGPA-----MQBALKLVLAEDGSALSRSKVLYLQVQL 227	Db
Qy	777 NENPVSSGRHGRTRKHYRKs-VRSRSHETKSRPLKLRILUNVSNKGDRVVECQLBTHN 835	Db	720 PQVLPQVSGKOSTGQVSYAQPAPVAVAQPTQPTLASSVDAHS-----EEFRSYEARREHD 281	

Qy	341	TGDPANLRLRPHLELLANID-----PNDPAVSPTEQLENAMVAKTV	384		Db	1879	LKPSSPSSDNL	1888
Db	895	EGLQG-----LESIGKDDYGFSGSQKLGEGFKOPIPASMP-QQIGPTPSLTQV	944					
Qy	385	VHGLVDFPQNYSRKGHETPQPQPN-----SKYKTSMCRDLRQCGCPRGTNCRPAH	436	RESULT 11	US-09-854-856-14			
Db	945	VH-----SAGRFRIVSPVPSRILRESKVFSEITDTVASTAQSFG-----MNLSHS	991		Sequence 14, Application US/09854856			
Qy	417	QEBLEKYR----LRNKKI-----NATVRTEFP-----LLNKVGVNNTVTPTA---	473		Patent No. 6541252			
Db	912	ASSLSLOCASEURRAQNTIEGPNTAPPNFSHTGPTFPVPPFLSS1AGVPTTAATAAPVP	1051		GENERAL INFORMATION:			
Qy	474	-----GNVISVIGSTETTGKVPSNGISNAENS-----VSOLIS	508		APPLICANT: Walke, D. Wade			
Db	1052	ATSSPPNIDISTSYQSET-----VPTERGDIAGVATSTGVVTSGGLPBPVSEPVLSVS	1108		APPLICANT: Donoho, Erin			
Qy	519	RSTDSLALAEVTKVGKVGAN-----GQNAGAPSAD-SVTEENKG-----SPPKTPVSN	557		APPLICANT: Turner, C. Alexander Jr.			
Db	1109	SITIPAVVISITSPSLQVPTSTSEIVVSSATLYPSVTSASASAGGSTATGPGRPPAV	1168		TITLE OF INVENTION: Human Kinases and Polynucleotides			
Qy	558	VAATSAGPSNVGTELNSV-----PKSSPFLTRVPPVPPHSENTQYF	599		TITLE OF INVENTION: Encoding the Same			
Db	1169	VSOQAAGGTTVGATLTSSTTTSFPSTATSOLSTQLOSSSTSTPLAETVVSAHSLD-KTS	1227		FILE REFERENCE: LEX-0178-USA			
Qy	600	QDRTQIIPPEVQPYQPGYYPBP-----TVPRAGVAFCP-PFRVRSNNVPESSLPPASMP	653		CURRENT APPLICATION NUMBER: US/09/854,856			
Db	1228	HSSSTGFLSF-SAPSSSSSPGAGVSSYTSOPGLHPVPISTASTPLPDAAGPTSTP	1286		CURRENT FILING DATE: 2001-05-14			
Qy	654	YADHYSTSFSPRDRMNSSYQPPPQPYEVPPVPPSGM-----YAPYDS	697		PRIOR APPLICATION NUMBER: US 60/206,015			
Db	1287	L-----LPQVHSIPLVQPVANPVAVQQTLLHSQOPALLPNQOPHTHCPEVDS	1334		PRIOR FILING DATE: 2000-05-19			
Qy	698	RRIWRPPMYORDI-----IRS--NSLPPMDVMHSS-YOTSLRERYNSLDGYSSA-	746		NUMBER OF SEQ ID NOS: 64			
Db	1335	DT-QPKAGIDDIKTLBKLRLSEIISSSGAQHASVLETSVLTGIPITAVA	1392		SOFTWARE: Fast-SEQ for Windows Version 4.0			
Qy	747	-----COPPSE PRRTYPLPRTC-----GHLKTSCEBOIRR KPDQWAQYHTQKA	790		SEQ ID NO: 14			
Db	1393	PSKULUTSTSTCPLPPTNPLGTVLVALPVPTVPTVPTVPTVPTVPTVPTVPTVPTV	1448		LENGTH: 2136			
Qy	791	PLVSS-TLIVPATQSPTPPSVDFRADSFSVSGTKFEEDHLSHYSPWSCOTIGCIN	849		TYPE: PRT			
Db	1449	PLTKAPVPLPVGTELP-----AGTLPS-E	1470		ORGANISM: Homo sapiens			
Qy	850	ATNSEPKDVIANSMVLMIDLS-----GDIVKRVHLFETORRTK---	888		FEATURE:			
Db	1471	QLPFFPGRPLTQSCQPLBLDAQLRRRLJSPEXITVTSVAVGSPMSMAAPTAITEGTOPQKG	1530		NAME/KEY: VARIANT			
Qy	889	-----EEDPPIPFSDGPISKWG-----AIRSSRRTGHHTDBQATASOGSATKPIVS	938		LOCATION: (1)...(2136)			
Db	1531	VSOYKEGPVLAATSSGAGVFMGRFQVSAADGQKEGNKSEDSVSESSVLS	1590		OTHER INFORMATION: Xaa = Any Amino Acid			
Qy	939	DYPPVNAV-----DSSWSSYGEANTS-SAHYVERDRIFTDLS-G	977		US-09-854-856-14			
Db	1591	SSSPESTLVKEPNGITIDGISSDVPESAHKTTASEASDTGQPTKVGFOVTTANKVG	1650		Length: 2136;			
Qy	978	HRKHSSTGDLS-----LELOOAKSNLLLOREANALAMQKWN-----1016			Best Local Similarity: 19.2%;			
Db	1651	RFSVSKTEDKITDKKEGPVAVSPFMDLEQAVLPAPIKEKPELSESHLNQSPSDEA	1710		Pred. No. 8.6e-16;			
Qy	1017	-----SLDEGR---MLTANLSSKBIERN-----GELQSDY-TEDATDKPORDIEEL	1061		Matches 256; Conservative 168; MinMatches 527;			
Db	1711	AFLSRDWDGSGSPHSIPHQSKLPSQNLQSLNSNSMSSDNNSDIEBDLKL	1770		Indels 379; Gaps 57;			
Qy	1062	SALDTDEPDQCGSPEEEILDQIOLGTSQNSNSMSSDNNSDIEBDLKL	1770					
Db	1771	RRL-----RDKLKEIQLQSRQKHESLTYKLKGVPVAVIPTPAAPLSGRRRRTK	1823					
Qy	1115	QK-----KQSLGEDH-----VILEEQKTLIPVTSCEFSQOLPVSVSNASCILP	1155					
Db	1824	SKGSKSSRSSLGNKSPQLQSGNLQGSQAAASVLHFPQTTHPPGN-----IPESCGNQLIQP	1878					
Qy	1156	ITTSVSGNL 1165						
Qy	1052	ASSLSLOCASEURRAQNTIEGPNTAPPNFSHTGPTFPVPPFLSS1AGVPTTAATAAPVP	1111					
Qy	1055	VH-----SAGRFITSPVPSRLESKVPESEITDTVAASTAQSPG-----MNLSHS	1051					
Qy	1057	OEELEKYR-----LRNKKI-----NATVTRTFP-----LNKVGVNNTVTATA	473					
Db	1058	-----GNDVVISGTTETGKIVPSTNGISNAENS-----VSOLIS	508					
Qy	1064	-----GQIGIPTSSLTQV-----LESIQLGKDDYGFSGSQKLGEGFKOPIPASMP	1004					
Db	1112	ATSSPPNDISTSV1QSBV-----VPTEEIAGVATSTGVVTSGGLPIPVSSESPVLSVVS	1168					
Qy	509	RSTDSTLRALEYTKVGKVGAN-----GNAJAGPSAD-SVTEENKG-----SPPKTPVSN	557					
Db	1169	SITIPAVVISITSPSLQVPTSTSEIVVSSSTALYPSVTSASASAGGSTATGPKPVAV	1228					
Qy	558	VAATSGPSNVGTELNSV-----PQKSSPFLTRVPPVPHSENIOFY	599					

; SOFTWARE: FastSEQ for Windows Version 4.0 ; SEQ ID NO 64 ; LENGTH: 1911 ; TYPE: PRT ; ORGANISM: Homo sapiens ; FEATURE: NAME/KEY: VARIANT ; LOCATION: (1) ..(1911) ; OTHER INFORMATION: Xaa = Any Amino Acid ; US - 09 - 854 - 856 - 64

Query Match		Score 193; DB 4; Length 1911;	
Best Local Similarity 19.3%; Pred. No. 7.8e-06;		Mismatches 159; Indels 378; Gaps 56;	
Matches 248; Conservative 159;			
Qy	698 RRIWRPPMQRDDI-----IRS--NSLPPMDYMHSSY-YQTSLRERVNSLDGYSVA--746	Qy	160 ERTVTEILILQHONPQLSANLWAARARGCQFLGPA---MQEZAHLKVLIALEDGSAL 214
Db	1395 DT--QPKAPGIDDIKTLEBEKRSLSFSEHSSGRQHAHSVLSLETSVTFPIPTAVAA 1452	Db	679 EATTAQPVSQPQAPQVLPQVSAGKUSTOGVSQVAPQDQPTLASSVDSAH 738
Qy	747 -----COPPSE PRITVPLPPEC --GHLKTSCEBQIRR-KPDQMAQHTQKA 790	Qy	215 SRKVLYLFVVORLEPREPOASAKSTISGHVWOLLYRASCFKYTKRDSDSILMOKL---E 268
Db	1453 PSKLTSSTSTCUPTNPLGTVLVALPVYVTPFGQVSTPVSTTSGVKPGT---APSXP 1508	Db	739 SD---VASGMDGNENVPSSSGRHGRCTTRHYRS-VRSRSRHKTSRKLRINVNNSK 794
Qy	791 PLVSS-TLUPVATOSPTPSPLSVDFRADSSESVSGTKFEEDHLSHYSWNSCGTIGSCIN 849	Qy	269 EFRSYEARRREHDQAQIVHIAEAGLRLISPCKOWSLLY-GDLAKHSHMOSTIDKLGQSPEF 327
Db	1519 PLTKAPVTPVGELP-----AGTLIS-E 1530	Db	795 GDRVVCQLETHNRKMYTFKFDPD-GDNPERIATMVNNDELAERESPVDQYREIEK 853
Qy	850 AIDSEPCKVIANNAVIMLDLS-----GPDVKRRVHLFETORRKT---888	Qy	328 AKSVQBLTIVLQRTGPANLNRLRPHLELLANID-----PNPDAVSPTW 371
Db	1531 QLPFPFGSLTQSQQPLFDLDAQLRRTLPVGELP-----AEGTQFQKG 1590	Db	854 ADEMILSEDVSPEZPEGDQG-----LESLGKDDYGFSGSSQKLEGEFKQPIPASSMP---903
Qy	889 ---EEDIIDIPFSDGPIISKWG-----AISRSRRTQHTIDPVQTAQSATRPIVS 938	Qy	372 EQLENAMEAVKTVVHGLVDFDIONYSRKGHETPQPOPEN-----SKYKTSMCRDLRQQG 423
Db	1591 VSQVKEGVYLATESGAGYFKMGRFOVSAADGAQKESKNSDADKSVHFEESTSESSVL 1650	Db	904 QKIGITPSSLTQVNH-----SAGRERFIYSPVPSRSLRESKVFSEITDTVAASTAQSP 956
Qy	939 DYPVYNNV-----DSRWSSYGEATS-SAHYVERDRFTVTDLS--G 977	Qy	424 GCPRGFNCNTFAHSSOEELEYKR---LRNKKI-----NATVTRTFP---LLNKV 463
Db	1651 SSSPESTIVKPEPENGITIPGISSDVPSAHTTASEAKSDTQOPTVGRFQVTTANKVG 1710	Db	957 G-----MNLHSASSLSLOAFSELRAONTEGNTAPPNFSHGTPFPVVPPLUSSIA 1010
Qy	978 HRRHSSSTGDL-----LEILOQAKNSNLJLQREANALAMQOKWN-----1016	Qy	464 GVNTNTVITA-----GNNVISIGSTETGKIVPSTNGISNAENS-----502
Db	1711 RFSVSKTADKDTKKEGPVASPFFMDLEQAVLPKKEPKPELSEPSHLNGPSDDPEA 1770	Db	1011 GVPTTAATAPVATSPSPNDISTSVIOSEVT-----VPTEEGIAVATSTGVTSGLPIP 1067
Qy	1017 ----SLEBGR--HUTINLISKEIERN----GEIQLSDY TEDATDKPRDIEEL 1061	Qy	503 -----VSQLISSRSDSTIRALETVRKVGKVGAN-----GQNAGAPSAD-SVTENKIG 548
Db	1771 AFLSRDVEGGSGSPHSPIQOLSSKLPSQNLSSNSFNNSSTMSSDNEDEDLKL EL 1830	Db	1068 PVSESPVLTSSVYSSSTIPAVVSISSTSPLSLOVPTSTSEIVVSSSTLYPSVTSASASAG 1127
Qy	1042 SALTDDEDGQGSPPIEFLDIQIGISSNDQLLN-----GMAVENGHPVQHQKEPK 1114	Qy	549 -----SPPKTPVSNVAAATSGPSNVYCTELNSV-----PQKSSPFLTRV 586
Db	1811 RRL-----RDGHKEITQDLSQRKEHIESLYTKLGKPPAVIIAPAISGRRRRPTK 1883	Db	1128 GSTATGPCKPAVVSQQAAGSTVQATLTSVTSSTSTLAET 1187
Qy	1115 QK-----KOSLGEDH-----VILBEOKTILLPVTSCEFSQLPVSISNASCLP 1155	Qy	587 PVYPPISENIOYQDQBRTOPIPEVQYQPGQYPPP-----TVAGVAFCV-PFVRVRN 640
Db	1884 SKCSKSRSRSSLUNKSPOLSGNLSGQZAAVSLHPQQTLHPPGN-----IPESQGNOLLQP 1938	Db	1188 VVVAHSLD-KTSHSSTTGIAFLS-SAPPSSSSPAGVSSYIS:OPQGLHLPLVPIVIAST 1245
Qy	1156 ITTSVAGNL 1165	Qy	641 NVPESSLUPPAMPYADHYSFTSPRDMNSSPYQPPPQPYGPVPPVPSGM-----690
Db	1919 LKPSPPSNL 1948	Db	1246 PILPQAGP1STPL-----LPQVPSIPLVQPYANPVAQVQTLIHSQPOPALL 1293
Qy	RESULT 12	Qy	691 -----YAPYDSRIRIWRPPMNYQRDDI-----IRS--NSLPPMDYMHSSV-YQTSRLR 735
Qy	US-09-854-856-64 Application US/09854856	Db	1294 PNQPHTHCPEVDSDT--QPKAGPIDIITLEEKLSLFSHESSSSGAQHASVSLSVIE 1351
i	Patent No. 6541252	Qy	736 YNSLIGYYSYA-----COPPKTPVPLPREFP--GHIKTSCEBQIRR-777
i	GENERAL INFORMATION	Db	1352 STVTPGTPAVASPKSLTSTCUPTMPLGTVALPPTPVWPGQVSTPVSTTSGV 1411
i	APPLICANT: Donoho, Gregory	Qy	778 KPDQWQYHTOKAPLYSS-TLPVATOSPTPPSPLSVDRADFESSVSGTKFEEBDLHSY 836
i	APPLICANT: Turner, C. Alexander Jr.	Db	1412 KPGT----APSKPKPLTKAPVLPVGTEL-----14356
i	TITLE OF INVENTION: Human Kinases and Polynucleotides	Qy	837 SPWSCTGIGSCINAIDSEPKDVIANSNAVIMLDLS-----GDKRVRVH 879
i	FILE REFERENCE: LEA-0178-USA		
i	CURRENT APPLICATION NUMBER: US/09/854,856		
i	PRIORITY FILING DATE: 2001-05-14		
i	PRIOR FILING DATE: 2000-05-19		
i	NUMBER OF SEQ ID NOS: 64		

US-09-854-856-32  
 Sequence 32, Application US/09854856  
 Patent No. 6541252  
 GENERAL INFORMATION:  
 APPLICANT: Walke, D. Wade  
 APPLICANT: Hilbun, Erin  
 APPLICANT: Donoho, Gregory  
 APPLICANT: Turner, C. Alexander Jr.  
 TITLE OF INVENTION: No. 6541252 el Human Kinases and Polynucleotides  
 FILE REFERENCE: LEX-0172A-USA  
 CURRENT APPLICATION NUMBER: US/09/854, 856  
 CURRENT FILING DATE: 2001-05-14  
 PRIOR APPLICATION NUMBER: US 60/206, 015  
 PRIOR FILING DATE: 2000-05-19  
 NUMBER OF SEQ ID NOs: 64  
 SOFTWARE: FastSEQ For Windows Version 4.0  
 SEQ ID NO 32  
 LENGTH: 1971  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: VARIANT  
 NAME/KEY: LOCATION: (1): (1971)  
 OTHER INFORMATION: Xaa = Any Amino Acid

US-09-854-856-32  
 Query Match 3.1%; Score 193; DB 4; Length 1971;  
 Best Local Similarity 19.3%; Pred. No. 8.2e-06;  
 Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

1436 ---AGTLP---EQLPPPGPSLTQSQOPPLEDAQRRLTLSPEXTTVTSAVGPVSMAAP 1489  
 810 LFETQRRTK-----EDPFLIPFSQPIIPII-----AISRSRRTGHTTDPVQAT 925  
 1430 TAITEAGTOPQKGVSQVEGPILATSSAGVFKMGRITQVSVAADGAQEGKNSKEDAKSV 1549  
 926 ASQGSATPKISYSDVYPPVNAV-----DSRWSYGNNEATS-SAHYER 967  
 1550 HPSSTSASSSSSSSPSTLKVPEPNHITPISSDPESAHKTTASEAKSDTGTQPTV 1609  
 958 DRFIVTDLI-----GHRKHSSTGDLI-----LELOQAKNSNLLQREANALAM 1011  
 1610 GRQVTTANKVGCRFSVSKTEBKDTKKEGPVASPPFDLQEAVLPVKEKPELSE 1669  
 1012 QOKWN-----SLDGR-----HHTLNILSKEBIELRN-----GELQSDY-TEDA 1048  
 1670 PSHLNGPSSDPEAFLSDEDDGSPSPHOLQSKIPSQNLQSQLSNSFNNSYMSSDN 1729  
 1049 TDTKPDRIELELSALDDEPDQOSEPEBEIIDLQLGSSONDOLIN-----GMAVEN 1101  
 1730 BSDIEDEKLKEIIRR-----RDKHKEIOPLSQRKHETESLYTKLGKVPATTIP 1782  
 1132 GHVHQHKEKEPKOK-----KOSLGEDH-----VILEEQKTILPVTSCFSQ 1142  
 1783 AAPSGRRRPTSKGSRSSSLGNKSPQLSGNLSCQSAASVLHQTLHPPGN--- 1838  
 1143 PLPVTSINASCLPITTYSAGNL 1165  
 1839 -IPESGQNQNLQPLKPSLSSDNL 1860

RESULT 13  
 Sequence 32, Application US/09854856  
 GENERAL INFORMATION:  
 APPLICANT: Walke, D. Wade  
 APPLICANT: Hilbun, Erin  
 APPLICANT: Donoho, Gregory  
 APPLICANT: Turner, C. Alexander Jr.  
 TITLE OF INVENTION: No. 6541252 el Human Kinases and Polynucleotides  
 FILE REFERENCE: LEX-0172A-USA  
 CURRENT APPLICATION NUMBER: US/09/854, 856  
 CURRENT FILING DATE: 2001-05-14  
 PRIOR APPLICATION NUMBER: US 60/206, 015  
 PRIOR FILING DATE: 2000-05-19  
 NUMBER OF SEQ ID NOs: 64  
 SOFTWARE: FastSEQ For Windows Version 4.0  
 SEQ ID NO 32  
 LENGTH: 1971  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE: VARIANT  
 NAME/KEY: LOCATION: (1): (1971)  
 OTHER INFORMATION: Xaa = Any Amino Acid

US-09-854-856-32  
 Query Match 3.1%; Score 193; DB 4; Length 1971;  
 Best Local Similarity 19.3%; Pred. No. 8.2e-06;  
 Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

160 ERTYELLIQHONPOQLSANLWAVRGCQFLGPA-----MQEEALKLVJIALEDGSAL 214  
 739 EATTAQPVQPOQAQVLPQVSAGKOSTQVSOVAPAEPVAVAQFOATOPTTASSVDASH 798  
 215 SRKVYLVLPVYQRLRERPFQASKTSIGHVVQLLVRASCVKTRDEDSIMQLK-----B 268  
 799 SD---VASGMSDGENVNPSSSGRHEGRTKHYRK-----VILEEQKTILPVTSCFSQ 1142  
 269 EFRSYEARREHDQAQIVHITAMEAGURISPSQWSSLILY-GDLAHKSHMOSIIDKLQSPPSF 327

Db 855 GDRVVVECQLETHNRKMVTKFKDLD-GDNPERIATIMVNNDPLATERESPVDQVRREIEK 913  
 Qy 328 AKSVQBLTIVLQRTGPANLNRLRPHELLAND-----PNPDAVSPTW 371  
 Db 914 ADEMSEEDSVBEPGEBOQ-----LESLGKDKDGYGFSGSQLKEGEFKQTPASSMP-- 963  
 Qy 372 EOLENAMAVKTVVHGLVDFQIONYSRKGKHETPQPQPN-----SKYKTMCRDILRQG 423  
 Db 964 QOIGIPTSSLQVWH-----SAGERPIVSPVPSRLRESKVPSSEITDVAASPAQSP 1016  
 Qy 424 GCPRTGTCFAHQSQEELEKYR-----LRRKKI-----NATVRTFP----LLNKV 463  
 Db 1017 G-----MNLSHSASSLSQLQAFSELRAQMTEGPNATTAPPNFSHTGPTFPVPPFLSSIA 1070  
 Qy 464 GVNNNTVTTA-----GNVISVGSSTETGKIVPSTNGISNAENS-----502  
 Db 1071 GYPTTAATAAPATSSPNDISTSVQSET-----VPEEGIAGVATSTGVUTSGGJLPIP 1127  
 Qy 503 -----VSQLIISRSTDSTALETYTKVGKRVGAN-----GQNAAGPSAD-SVTEENKIG 548  
 Db 1128 PVSESPVLSVSVSITIPAVVSISSTPSLQVPTSEIVVSSTALYPSVTVSASAG 1187  
 Qy 549 -----SPPKTPYNSVAAATSAGPSNVGTELNSV-----PKQSSPFLTRV 586  
 Db 1188 GSTATPGPKPAAVSSQQAAGSITVGAATLTSSTTSPPSTASSQLTOSLSSSTSTPLAET 1247  
 Qy 587 PVYPPHSENIVQFQDPPTQIPIPEPVQYGPQYGPYPPP-----TVAGVAPCV-PRFVRSN 640  
 Db 1248 VVVAHSULD-KTISHSSTTGIAFLSL-SAPSSESSSSPAGVSSYISQPGSLUHPLVTPSVIAST 1305  
 Qy 641 NYPESSLPPASMPYADHYSTFSPRDRMNNSPQPPFPQYGPVPPVSGM-----690  
 Db 1306 PILPQAAGPTSTPL-----LPQVPSIPLVQPVANPVAVQOTLHSQPQPAL 1353  
 Qy 691 -----YAPVYDSRIRWPPMYORDI-----IRS-----NSLPPMDYMHSV-YOTSLRER 735  
 Db 1354 PNQPHTHCPEVDSDT-----QPKAFGIDDIKTLERKLRLSPFSEHSSGAGHSVLETSSIV 1411  
 Qy 736 YNSLDGYSVA-----COPPSE-PRTVPLPREC-- -GHLKTSCEQIRR- 777  
 Db 1412 SRTVPGIPTTAVAPSKLTSTSTCLPPTNLPLGTVALPVTPVTPVSTPVSTRTSGV 1471  
 Qy 778 KPDQWAQHTQKAPLYSS-TLPVATQSPTPPSPLFSVDFRADFSESSVGTKEBEDLHSY 836  
 Db 1472 KPT-----APSKPPLTKAPVPLVGTEP-----1495  
 Db 837 SPWNSCGTIGSCINAIDSEPKDVIANSAVNLMDLS-----GDVKRRVH 879  
 Db 1496 -----AGTLPS-----BQLPFPGPSPLTQSCQPLEDIAQLRRTLSPEXITTVTSAVGPVSNMAAP 1549  
 Qy 880 LFPTTQRTK-----BEDPLIPFSGPILLISKWG-----AIRSSRTGHTTDPMQAT 925  
 Db 1550 TAITEAGTOPQKGVSQYKEPVLATSSGAGVKMGRFQVSAADQKEGKNSKEDAKSV 1609  
 Qy 926 ASQSATPKISYSDYVPPYNAV-----DSRWSYGEANTS-SAHTVER 967  
 Db 1610 HFESTSSVTSUSSPSPSTLVKEPEPNGITIPGISSDVPESAHKTTASEAKSDTQPTKV 1669  
 Qy 968 DRPIVTDLS-----GHRKHSSTGDLI-----LELOQAKNSNLLQREANALAM 1011  
 Db 1670 GRQVTTANKVGRFSVSKTEKIDTKKEGPVASPPFMDLQAVLPVKKPPELSE 1729  
 Qy 1012 QQKWN-----SLDGR-----HLTNLILSKETELRN-----GELSDY-TEDA 1048  
 Db 1730 PSHNGPSSDPEAAFLSRDVDDGSGSPSPHQLSKSLPSQNLISQSLNSFNNSYMSDDN 1789  
 Qy 1049 TDTPKDRDIEELSALDTPDCQSEPIEELDOLQGSSQNDQLN-----GMAVBN 1101  
 Db 1790 ESDIEDEDUKLERL-----RDHKIKEIQQLQSRQKHLESLYTKLGKVPAPIPP 1842  
 Qy 1102 GHVQOHQKEPPKOK-----KOSLGEDH-----VILEEQKTILPVTSCFSQ 1142  
 Db 1843 AAPISGRRRPTKSQSKSSRSSSSLGNSKSPOLSGNLSCQSAASVLPQQTLPHPGN--- 1898

Qy 1143 PLPVSIASCLPLITTSVSGNL 1165  
 Db 1899 -IPESGQNQLLQFLKPSFSDNL 1920

**RESULT 14**  
 US-09-854-856-62 Application US/09854856  
 Patent No. 6541252  
 GENERAL INFORMATION  
 APPLICANT: Walke, D. Wade  
 APPLICANT: Hilbun, Erin  
 APPLICANT: Donoho, Gregory  
 APPLICANT: Turner, C. Alexander Jr.  
 TITLE OF INVENTION No. 6541252e1 Human Kinases and Polynucleotides  
 TITLE OF INVENTION Encoding the Same  
 CURRENT APPLICATION NUMBER: US/09/854.856  
 FILE FILING DATE: 2001-05-14  
 PRIOR APPLICATION NUMBER: US 60/206,015  
 PRIOR FILING DATE: 2000-05-19  
 NUMBER OF SEQ ID NOS: 64  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 62  
 LENGTH: 2048  
 TYPE: PT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: VARIANT  
 LOCATION: (1) ..(2048)  
 OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-854-856-62

Query Match Score 193; DB 4; Length 2048;  
 Best Local Similarity 19.3%; Pred. No. 8. 8e-06;  
 Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

Qy 160 ERTVTELJLQHNPQQLSANILWAAVRARGCQFLGPAA----MQEFAKLVTLALEGSAL 214  
 Db 679 EATTAQPSQSPQAQPQVNLQVSAGKQSTGVSQVAPAPSPVAYAQPQTQPTTASSYDSAH 738

Qy 215 SRKVVLVEFLVQRLEPRFFQAKSTSIGHVQLLYRASEKFVTKRDEDSLMLK---E 268  
 Db 739 SD---VASGMDGNENYVESSGRHEGRTRKRYRKVSRSRHEKTSRPKLRLNNVK 794

Qy 269 EPPSYEARREHDAQIVIYAMEAGLRTPEQWNSLLY-GDLAHKSHMQSIIIDKLQSPESP 327  
 Db 795 GDRVVECCOLETNKRMYTFKFOLD-GDNPSEEATIMVNDFIAERESFDQVREIEK 853

Qy 328 AKSVQELTIVLQRTGDPPLNLRLRPHLELLANID-----PNNPDAVSPTW 371  
 Db 854 ADEMNLSEEVSVPEGDQG-----LESLGKDDYRPHLELLANID----- 903

Qy 372 EQLENAWAVAKTVVHGLDFIQNYSRKHGTHETPOPQNP-----SKYKTSMCRDLRQG 423  
 Db 904 QQGIGPTEPLSTQVHV-----SAGRFIVSPVSPVSPVSPVSPVSPVSPVSPVSP 956

Qy 424 GCPRGNTNFARHQQEELEKIR---LENKKI-----NATVRTFP-----LINKV 463  
 Db 957 G-----MNLSHSASSLSQLQAFSELRLRAQMTEGPNTPAPPNSHTGTFPVVPPPELSSIA 1010

Qy 464 GVNNTVTTA-----GNYTISVGSTETTGKIVPSTNGISNAENS----- 502  
 Db 1011 GVPTTAATAPVDPATSSPNDISTSVIQSEVT---VPTEEGLAGVATSTGVVTSGGLPIP 1067

Qy 503 -----VSOLISRSTSTDSTLRALETVKVKGKYGAN----CQNAAGPSAD-SVTENKIG 548  
 Db 1068 PVSESPVSPVSVSISITPAVVSISSTSFSLQOPTSTSEIVVSSATASVATASAG 1127

Qy 549 ----SPPTTPVSNVAAATSAGPSNNGTELNVS-----PKQSSPFLTRV 586  
 Db 1128 GSTATPGKPKPAVVSQQAGSTRVATLTSVSTTSSPSTAQSOLSIQSSSSTPTLAET 1187

Qy 587 PVYPHSENIOYQODPRTQIPFEPVPOYPQTGYPPPPP-----TVPAGVARCV-PFRVRSN 640  
 Db 1188 VVVAISLSD-KTSHSSTTGFLFSL-SAPSSESSSPAGVSSYISQDQGLHPVIVPIVIAST 1245

Qy 641 NVPESSUPPASMPYADHYSTFSPRDMNSSPYQPPPQPYGPVPPVPPSGM----- 690  
 Db 1246 PILPOAAGPTSTPL-----LPQVSPSIPLVQPVANPVAVQOTLJHSQPOPALL 1293

Qy 691 -----YAPVYDSRSLWRPMYQRDDI-----IRS--NSLPPMDMHSY-YQTSRLR 735  
 Db 1294 PNQPTHICPEVDSDT--QPKAGIDDIKTLKLEKSLFSEHSSGAQHASVSELSVIE 1351

Qy 736 YNSLDGYSVA-----COPSPVPLPREPC--GHLKTSCEBQIRR- 777

Qy 778 KPDQWQAQHTQKAPLYSS-TLPVATQSPTPSPLTPTVQVSTPVSTSTSGV 1411

Db 1352 STVTPGIPTAVAPSKULLSTSTCUTPNTPLGTVALPPTPVVTPVQVSTPVSTTSV 1436

Qy 1412 KPGT---APSKPKPLTKAPVLPVGTELP----- 1435

Db 1432 KPGT---APSKPKPLTKAPVLPVGTELP----- 1435

Qy 1433 SPWSCGTIGSCINAIUSEPQDVIAINSAVNLMDLS----- 1436

Db 1436 ---AGTLPs--EQLPPFPGPSTQSLQQQPLEDLDQALRRTLSPEXKITVTSAVGPVSMAAP 1489

Qy 880 LFETQRTK-----EEDP11PFEDGPT1ISKWG-----AISRSSRTGHTTDPVQAT 925

Db 1490 TAITEACTQPKGVSVQVKGPVLAASSGAGYFKMGRFOVSAADEAQKEGKNSDAKSV 1549

Qy 926 ASQGSATKPLKPSVSDVPPVNAV----- 926

Db 1550 HFESETSESSVLSLSSSPESTLVKPEPNGIT1PGISSLVPESAHKTASEAKSDIGOPTKV 1609

Qy 968 DRPIVTDLS---GHRKHSSTCDLLS----- 968

Db 1610 GRFQTTTANKVGRPSVKTEDKTTDKKEGPVASPSPFMDELQATLPAVPKKEKPELSE 1669

Qy 1012 QOKWN-----SLDBEGR---HLTUNLSSKBEIERN-----GELQSDY-TEDA 1048

Db 1670 PSHLNGPSSDDEAFLUSRDVDDGSSPSPHQLOSSKSLPSNOLSSLSFNNSYSSDN 1729

Qy 1049 TDTKPDIELSALDTDEPDGOSSEPIEELDOLGISSONDOLLN-----GMAVEN 1101

Db 1730 ESDIEEDLKLRLR-----RDKHLKEIQLDLSRQKHEIESLYTKLGKUPPAVIPP 1782

Qy 1102 GHPVQOHQKEPKOK-----KOS1GEDH-----VIEBQKTLIPVTSFCFSQ 1142

Db 1783 AAPLGRERRRPTKSQSKSSRSSSJKNSPQLSGNILSGQSAASVHLHQQTULHPPGN---- 1838

Qy 1143 PLPVTSNASACLUPTSVSAGNL 1165

Db 1839 -IPESQONQLIQPLKSPSSDNL 1860

RESULT 15  
 US-09-854-856-30  
 Sequence 30, Application US/09854856  
 Patent No. 6541252  
 GENERAL INFORMATION:  
 APPLICANT: Walk, D. Wade  
 APPLICANT: Hilbun, Erin  
 APPLICANT: Donoho, Gregory  
 APPLICANT: Turner, C. Alexander Jr.  
 TITLE OF INVENTION: No. 6541252 Human Kinases and Polynucleotides  
 CURRENT APPLICATION NUMBER: US/09/854.856  
 CURRENT FILING DATE: 2001-05-14  
 PRIORITY NUMBER: US 60/206,015  
 PRIORITY FILING DATE: 2000-05-19  
 NUMBER OF SEQ ID NOS: 64  
 SOFTWARE: FastSEQ for Windows Version 4.0  
 SEQ ID NO 30

LENGTH: 2108  
 ; TYPE: PROTEIN  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1) ..(2108)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-854-856-30

Query Match 3.1% Score 193; DB: 4; Length 2108;  
 Best Local Similarity 19.3% Pred. No. 9.2e-06;  
 Matches 248; Conservative 159; Mismatches 498; Indels 378; Gaps 56;

Qy 160 ERTVTELLQHONPQLSANLWAVRARGCQFLGPA----MQEBAKLVLLAEDGSA 214  
 Db 739 EATTAQPSQPQAVLQEVSAKGKQSTGVSQVAPAEPAVQAPOAQOPTTLASSVD SAH 798  
 Qy 2115 SRKVLVLFVQRLLEPRPQASKTSIGHYVQLIYTRASCFKVTRDEDSLMQLK---E 268  
 Db 799 SD---VASGMSDGNENVSSSGRHEGRITKRHYRKs-VRSRSRHEKTSRKURILAVSNK 854  
 Qy 2669 EPFSYEYEARREHDAQIVTIAAMEGLRISPEQNSLLYGDLAHKSHNMQSIIIDKLQSPESF 327  
 Db 855 GDRVVEBCOLETHRNKMVTTFKFDL-GDNPEETATIMNDFLIAERESFDQVREIEK 913  
 Qy 328 AKSVQELTIVLQSTGDPANLNRLRPHLELLND----PMPDAVSPTW 371  
 Db 914 ADEMLSEDVSVEPGDQG----LESLOGKDDYGRSGSQKLEGFEKFQPPASSMP-- 963  
 Qy 372 EOLENAMYAVAKTVYHGLVDFIQNYSRKGHETPOPQPN-----SKYKTSMCRDLJrqQ 423  
 Db 964 QQGIPPTSLTQVNH----SAGRRTIVSPVPESRILRESKVFVFPSE-TDTVAASPAQSP 1016  
 Qy 424 GCPRGTNCTFAHOOEELEYK---LRNKKI-----NATVTFP----LINKV 463  
 Db 1017 G-----MNLSHSSASSLSQLQAFSELRLRAQMTEGPNTPAPPNFSHTGTPVVPFFLSSIA 1070  
 Qy 464 GVNNTVTIA-----GVNIVSIGSTETTKIVKPATNGISNAENS----- 502  
 Db 1071 GVPTAAATAPVATPSSPNDISTSVQSEVT---VPTEEGIAGVASTGVUTSGGLPIP 1127  
 Qy 503 -----VSQLISRSTDSTLRALETVKVGKVGAN----GONAAGFSAD-SVTENKIG 548  
 Db 1128 PVSESPVLSVSSSITIPAVSISSTSPLSQLVPTSTSIVSSTALYFSTVTSATSASAG 1187  
 Qy 549 -----SPPTPVSNVAAATSAGPSNNGTTELNSV-----POISSSPFLTRV 586  
 Db 1188 GSTATPGPKPAPVSYQQAQSTGTVGATLTSVSTTSFPSTAQSOLSIQLSSSTSTPTLAET 1247  
 Qy 587 PVPDPHSNTIQYFODPRTPIFEPVQPYOTGYVPPP----TVPAQVAPCV\_PRFVRSN 640  
 Db 1248 VVUSAHSLD\_KTSHSSTSGLAFLSL-SASSSSSSPGAGTSSYISQPGGHLVIPSVASt 1305  
 Qy 641 NVPESLPPASMPYADHYSTSFSPDRMNSPYOPPPPPGPVPPVPGM----- 690  
 Db 1306 PILQOAGPSTPL-----LpQYSP1PLVQEVANTPAVQOTLHISQPQALL 1353  
 Qy 691 -----YAPYDSRRIWRPMPYQBDI----IRS-NSLNPMDVMHSS\_YOTSLRER 735  
 Db 1354 PNQPHTHCPEVDSDT-QPKAPGDDIKTLEEKLRSLSLSEHSSSGAOHASVSLTSLVIE 1411  
 Qy 736 YNSLDGYYSA-----CQPPBEB-PRTVPLPREPC---GHLKTSCEBQIRR- 777  
 Db 1412 STVTPGIPTAVASKLILSTSTCLPPTNLPGTVAlPVTPTPGVQSTPVTSTTISGV 1471  
 Qy 778 KPDQWQAQHTOKAPLVS-TLPVATQSPTRPSPFLFSVNDRADFSESVCTKFBEDHLSHY 836  
 Db 1472 KPGT---APSKEPLKPAVPLPVCTELP----- 1495  
 Qy 837 SPWCGTIGSCINAIIDSEPKDVIANSAVLMIDS-----GDVKRKH 879  
 Db 1496 ----AGTLPs--EQLPFPFGPSLTSQSOQPLDEDLQARLRTLSPEXITYTSAVGPVSMMAp 1549

880 LFETQRRTK-----EEDP1IIPFSGDGP1IISKWG-----AIRSSRTGYHTTDVQAT 925  
 Db 1550 TAITEAGTQPKGVSVQKEPVLAESSGAGYFKMGFQVSYAADDQKEGKNKSDEAKSV 1609  
 Qy 926 ASQGSAKTPKISVSDTYPYNNAV-----DSRWSSYGNREATS-SAHYVER 967  
 Db 1610 HFESETSESSVYLSSSPESTLVKEPENGITIPGISSDVPEAHKMTASEAKSDTGOPTKV 1669  
 Qy 968 DRFIVTDLS---GHRKHSSTDLLS-----LELOQAKNSNLLQREANALAM 1011  
 Db 1670 GRFQVTTITANVGRPSVSKTEDKITDKKEGPVASPPFMDEQAVLPAVIKKEKEPBELSE 1729  
 Qy 1012 QQKWN-----SLDBGR---HLTUNLLSKKEIELN-----GEJQSDFY-TEDA 1048  
 Db 1730 PSHLINGPSSDPEAAFLSRDVGSGSPSPHOSKSLPSQNLSSLSNSFNSYMSDDN 1789  
 Qy 1049 TDTKPDIDIEFLSALDTDDEPDGOSERPIBEILDIGLGISSQNDOLLN-----GMAVEN 1101  
 Db 1790 ESDIEDDLKUELRU-----RDGHLEKEQDLQSRQKHIEIESLYTKLGKVPATLIPP 1842  
 Qy 1102 GHPVQOHQKEPPPKK-----KQSLGEDH-----VILEEQKTILPVTSCFSQ 1142  
 Db 1843 AAPLSGRERRRPTKSGSKSSRSSSSLGNSKSPQLSGNLSGGQSAASVLUHPQTULHPPGEN--- 1898  
 Qy 1143 PLPVVISNASACLPITSVSAGNL 1165  
 Db 1899 -IPESGNQLLQPLKSPSPSSDRL 1920  
 Search completed: August 4, 2005, 01:49:10  
 Job time : 56 secs

This Page Blank

This Page Blank (uspc)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: August 4, 2005, 01:27:38 ; Search time 192 Seconds  
(without alignments)

3176.491 Million cell updates/sec

Title: US-10-619-992-2  
Perfect score: 6215  
Sequence: 1 MPVQAAQWTEFLSCLSPICYNE.....MSEDKNDFLKPYANGKVN1 1191

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:

1612378

RESULT 1  
Q9HBD2

ID Q9HBD2; PRELIMINARY;  
AC Q9HBD2; DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update).  
DE Membrane-associated nucleic acid binding protein (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Buteraria; Primates; Catarrhini; Hominidae; Homo.  
RT "A human gene coding for a membrane-associated nucleic acid-binding  
protein.";  
RN [1];  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20517928; PubMed=0938876; DOI=10.1074/jbc.M004161200;  
RA Siess D.C., Veeder C.T., Merkens L.S., Tanaka T., Freed A.C.,  
RA McCoy S.L., Heinrich M.C., Deffebach M.E., Bennett R.M.,  
RA Hefeneider S.H.

Database : UniProt 03:  
1: uniprot\_sprot:  
2: uniprot\_trembl:  
Pred. No. is the number of results Predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5.83	88.2	1048	2 Q9HBD2	Q9hbd2 homo sapien
2	4.12	71.0	869	2 Q9NKE1	Q9nke1 homo sapien
3	3.721	59.9	728	2 Q8ND6	Q8nd6 homo sapien
4	2.847	45.8	540	2 Q9HSJ2	Q9hsj2 homo sapien
5	2.685	43.2	1136	2 Q69231	Q69231 mus musculus
6	2.577.5	41.5	1109	2 Q8IVT6	Q8ivt6 homo sapien
7	2.477	39.9	1114	2 Q6ND6	Q6nd6 xenopus lae
8	2.293.5	36.9	506	2 Q8EST6	Q8est6 homo sapien
9	2.165	34.8	419	2 Q9HBD1	Q9hbd1 homo sapien
10	1.558	24.7	819	2 Q9VY48	Q9vy48 drosophila
11	1.289.5	20.7	398	2 Q8PTD3	Q8ptd3 anopheles g
12	0.889	14.3	198	2 Q9NPN9	Q9npn9 homo sapien
13	0.843	13.6	1014	2 Q459E2	Q459e2 caenorhabdi
14	0.780	12.6	177	2 Q9FZ27	Q9fz27 homo sapien
15	0.611	9.8	522	2 Q8NVY1	Q8nv1 homo sapien
16	213.5	3.4	1157	1 BBC1_YERST	P47068 saccharomyces cerevisiae
17	2.21	3.4	5085	1 Q6CLO1	Q6clo1 rattus norvegicus
18	2.11	3.4	1918	2 Q752A6	Q752a6 ashbya gossypii
19	207.5	3.3	1090	2 Q9PSJ0	Q9psj0 neurospora crassa
20	207.5	3.3	1100	2 Q7R2C6	Q7r2c6 neurospora
21	205	3.3	1289	2 Q9FLQ7	Q9flq7 arabidopsis thaliana
22	206	3.3	1611	2 Q42854	Q42854 schizosaccharomyces pombe
23	205	3.3	1730	2 Q6FJN8	Q6fn8 candida glaucomanniana
24	204.5	3.3	9234	2 Q7KTP5	Q7ktp5 drosophila melanogaster
25	204	3.3	2377	1 WNK1_MOUSE	P8741 mus musculus
26	203	3.3	786	2 Q48809	Q48809 arabidopsis thaliana
27	202	3.3	760	2 Q9T0K5	Q9t0k5 arabidopsis thaliana
28	201	3.2	1082	2 Q8VIG1	Q8vig1 mus musculus
29	200.5	3.2	728	2 Q9HPN3	Q9hpn3 candida tropicana
30	200	3.2	1134	2 Q63ZWS	Q63zws mus musculus
31	198	3.2	684	2 Q6VQRO	Q6vqro homo sapien

Query Match 88.2%; Score 5483; DB 2; Length 1048;  
Best Local Similarity 100.0%; Pred. No. 1.7e-265;  
Matches 1048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPVQAAQWTEFLSCLSPICYNE.....MSEDKNDFLKPYANGKVN1 60  
Db 1 MPVQAAQWTEFLSCLSPICYNE.....MSEDKNDFLKPYANGKVN1 60

Qy 1 IDVLPVNFAFLQLQVQPDHOSIKLNGLNRKACPPDQTAINTD 60  
Db 1 IDVLPVNFAFLQLQVQPDHOSIKLNGLNRKACPPDQTAINTD 60

Qy 1 IDVLPVNFAFLQLQVQPDHOSIKLNGLNRKACPPDQTAINTD 60  
Db 1 IDVLPVNFAFLQLQVQPDHOSIKLNGLNRKACPPDQTAINTD 60

Qy 1 INQSAISRMQRKLVTLYNCOLVVEEGRVRAARSIGERTYTELILQHQNPQLSANSI 180  
Db 1 INQSAISRMQRKLVTLYNCOLVVEEGRVRAARSIGERTYTELILQHQNPQLSANSI 180

Qy 121 LNQSAISRMQRKLVTLYNCOLVVEEGRVRAARSIGERTYTELILQHQNPQLSANSI 180  
Db 121 LNQSAISRMQRKLVTLYNCOLVVEEGRVRAARSIGERTYTELILQHQNPQLSANSI 180

Qy	181	WAARARGCQFLGPAMOEBALKLVLIAEDGSALSRLKVLVPRFQASAKTSIG	240	OX	NCBI_TaxID=9606;
Db	181	WAARARGCQFLGPAMOEBALKLVLIAEDGSALSRLKVLVPRFQASAKTSIG	240	RN	SEQUENCE FROM N.A.
RA	kawakami T., Noguchi S., Itoh T., Shigeta K., Senna T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibano T., Tanaka T., Nakamura Y., Isogai T., Sugano S.; Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		RP		
RA			RA		
RL			RA		
DR			RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; HGNC:8212; BAA91073.1;		DR		
Gene			DR		
GO			DR		
Intero			DR		
Pfam			DR		
SMART			DR		
SQ	SEQUENCE	869 AA; —96347 MW;	SQ	28105761C60B3D65 CRC64;	
			Query Match	71.0%;	
			Best Local Similarity	99.6%;	
			Matches	Pred. No. 4.2e-212;	
			Conservative	2; Mismatches	
			Indels	0;	
			Gaps	0;	
Qy	301	SSLYGDAHKSHMQSIDKLQSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANI	360	Qy	196 MOEALKLVVLAEDGSALSRLKVLVPRFQASAKTSIGHVQLYRASCERKV
Qy	301	SSLYGDAHKSHMQSIDKLQSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANI	360	Db	255 MOEALKLVVLAEDGSALSRLKVLVPRFQASAKTSIGHVQLYRASCERKV
Db	301	SSLYGDAHKSHMQSIDKLQSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANI	360	Qy	256 KRDEDDSIMQIKEEFRSYEALRREHDAQIVHIAEGLRISPEQWSLLYCDLAKHSHMQ
Qy	361	DPNPDAVSPTEOLEMANAVAKTCVHGUDFIONYSRKGHETPQPQENSKYKTSMCRDLR	420	Db	315 KRDEDDSIMQIKEEFRSYEALRREHDAQIVHIAEGLRISPEQWSLLYCDLAKHSHMQ
Db	361	DPNPDAVSPTEOLEMANAVAKTCVHGUDFIONYSRKGHETPQPQENSKYKTSMCRDLR	420	Qy	61 KRDEDDSIMQIKEEFRSYEALRREHDAQIVHIAEGLRISPEQWSLLYCDLAKHSHMQ
Qy	421	QQGGCPRTNCPAHSQEELEYRLRUNKINATVTRPLINKVGVNNTTTAGNIVSI	480	Db	120 STIDKLOSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANIDNPDAVSPTEOLE
Db	421	QQGGCPRTNCPAHSQEELEYRLRUNKINATVTRPLINKVGVNNTTTAGNIVSI	480	Qy	316 STIDKLOSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANIDNPDAVSPTEOLE
Qy	481	GSTBTGKIVPSTNGISMAENSVSQSLRSRSTDSLRRAETVKVGKGANGONAAGPSAD	540	Db	121 STIDKLOSPESPAKSVOELTIVLQRTGDPANLNRLRPHLEIANIDNPDAVSPTEOLE
Db	481	GSTBTGKIVPSTNGISMAENSVSQSLRSRSTDSLRRAETVKVGKGANGONAAGPSAD	540	Qy	376 NAMAVAKTVVHGLVDF1QNSRSGKGETPQPQPNSYKXTSMCRDLRROGGCPRGNTCTFAH
Qy	541	SYTENKIGCSPPKTPVSNVAATASAGPSNYGTELNSVPORKSSPPLTRVYVPPHSENIOYFQ	600	Db	181 NAMAVAKTVVHGLVDF1QNSRSGKGETPQPQPNSYKXTSMCRDLRROGGCPRGNTCTFAH
Db	541	SYTENKIGCSPPKTPVSNVAATASAGPSNYGTELNSVPORKSSPPLTRVYVPPHSENIOYFQ	600	Qy	436 SOBELEYKRLRNKKINATVTRPLANKVGVNNTTTAGNIVSITGKIVPSTNG
Qy	601	DPTQTIPPEVPOQPOTGTYPPPTVPGACVPRFVRSNNVPESSLPAASMYPADHYST	660	Db	240 SOBELEYKRLRNKKINATVTRPLANKVGVNNTTTAGNIVSITGKIVPSTNG
Db	601	DPTQTIPPEVPOQPOTGTYPPPTVPGACVPRFVRSNNVPESSLPAASMYPADHYST	660	Qy	241 SOBELEYKRLRNKKINATVTRPLANKVGVNNTTTAGNIVSITGKIVPSTNG
Qy	661	FSPDRDNNSPYOPPPPOPYGVPPVPGMAYPYDSDRIRWPPMYYORDDIRNSLIPM	720	Db	300 ISNAENSYOLISRSSTDSTRLALETVKVGKGANGONAAGPSADSYENKIGSPPKTPV
Db	661	FSPDRDNNSPYOPPPPOPYGVPPVPGMAYPYDSDRIRWPPMYYORDDIRNSLIPM	720	Qy	496 ISNAENSYOLISRSSTDSTRLALETVKVGKGANGONAAGPSADSYENKIGSPPKTPV
Qy	721	DYKHSSVYQTSLRLERYNLSLDGTYSVACOPPSEPRTRITYLPRECQGHLKTSCEQIRRKD	780	Db	301 ISNAENSYOLISRSSTDSTRLALETVKVGKGANGONAAGPSADSYENKIGSPPKTPV
Db	721	DYKHSSVYQTSLRLERYNLSLDGTYSVACOPPSEPRTRITYLPRECQGHLKTSCEQIRRKD	780	Qy	556 SVAATSAGPSNVGTELNSVPOKSSPPLTRVYVPHSENIOYFQDPRTRQIPFEVPOYPO
Qy	781	QWAQHTQKAPLYSSTLPEVATOSPTPPSPPLSFVDRADFSVSYSGTKFEEDEHLSHYSPWS	840	Db	615 616 TGYYPPPPTVPGAVPCPVFRVRSNNVPESSLPAASMYPADHYSTSPRDMMNSPQYPP
Db	781	QWAQHTQKAPLYSSTLPEVATOSPTPPSPPLSFVDRADFSVSYSGTKFEEDEHLSHYSPWS	840	Qy	421 TGYYPPPPTVPGAVPCPVFRVRSNNVPESSLPAASMYPADHYSTSPRDMMNSPQYPP
Qy	841	C GTIGSCNAIDSEPDKVTDNSAVMLDSDPVKRVHLFETQRRTKEEDP1PFSDFGP	900	Db	676 PQQYGPVPPVPSGMAPVYDSDRRIRWPMDYRNSLPPDMWHSSYQTQSLRER
Db	841	C GTIGSCNAIDSEPDKVTDNSAVMLDSDPVKRVHLFETQRRTKEEDP1PFSDFGP	900	Qy	681 PQQYGPVPPVPSGMAPVYDSDRRIRWPMDYRNSLPPDMWHSSYQTQSLRER
Qy	901	IISRWGA1SRSSRTGYHTDPVQATASSGATKPIVSYYVYNAVDSRWSYGEATS	960	Db	736 YNSLDGYYSAQCPSEPRTRVPLPREPGHILKTSCEQIRRKPQDWAQHTQKAPLYSS
Db	901	IISRWGA1SRSSRTGYHTDPVQATASSGATKPIVSYYVYNAVDSRWSYGEATS	960	Qy	796 TLPVATOSPTPPSPLSVDFRADFSVSQGKTFEEDLHSYHSPWSCGTGSCINADSEP
Qy	961	SAHYVERDRFIVTDLSGSRKHSSTGDLIISLLEQAKASNLIQREANALAMQKWNSLDE	1020	Db	800 541 YNSLDGYYSAQCPSEPRTRVPLPREPGHILKTSCEQIRRKPQDWAQHTQKAPLYSS
Db	961	SAHYVERDRFIVTDLSGSRKHSSTGDLIISLLEQAKASNLIQREANALAMQKWNSLDE	1020	Qy	856 KDVIANENAVMLMDSDVCKRVLFTTORRKEEDP1PFDGPTISKWGAISRSRTG
Qy	1021	GRHLTLNLISKEELRNCGELQSDYTEDA	1048	Db	916 661 KDVIANENAVMLMDSDVCKRVLFTTORRKEEDP1PFDGPTISKWGAISRSSTG
Db	1021	GRHLTLNLISKEELRNCGELQSDYTEDA	1048	Qy	971 721 YHTTDPYQATASQGATKPIVSYYVNAVDSRWSYGEATSAAHYVERDRFIVTDL
Qy					
RESULTS 2					
QNXKE1					
ID	Q9NXKE1	PRELIMINARY;	PRT;	869 AA.	
AC	Q9NXKE1;				
DT	01-OCT-2000 (TREMBLrel.	15	Created)		
DT	01-OCT-2003 (TREMBLrel.	25	Last sequence update)		
DE	Hypothetical protein FLJ20301.		Last annotation update)		
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
OC					

Db	781	SGRKHSSGTGDLISLELQAKNSNLLQREANALAMQQKWNSLDEGRHTLNLSKTEL	840	Qy	817	ADPESVSGTKEFEDHLSHYSPMSCGTTGSCINAIDSEPKDVIANSNAVLMOLDSGDVKR	876	
Qy	1036	RNGBLQ	1041	Db	481	ADPESVSGTKEFEDHLSHYSPMSCGTTGSCINAIDSEPKDVIANSNAVLMOLDSGDVKR	540	
Db	841	RNGEVK	846	Qy	877	RVHFETORRTKEFDPIIPSDGPPIISKGWAISSRSGTGYHTDPVQATASQSATPIS	936	
<b>RESULT 3</b>								
Qn3D6	TD	OBIN3D6	PRELIMINARY;	PRT;	728	AA.		
AC	QBN3D6;							
DR	01-OCT-2002	(TREMBLrel.	22,	Created)				
DR	01-OCT-2002	(TREMBLrel.	22,	Last sequence update)				
DR	01-OCT-2003	(TREMBLrel.	25,	Last annotation update)				
DB	Hypothetical protein	DKFZp762N035	(Fragment).					
GN	Name=DKFZp762N035;							
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;							
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
OX	NCBI_TaxID=9606;							
RN	[1]							
RP	SEQUENCE FROM N.A.							
RC	TISSUE=Melanoma;							
RG	The German cDNA Consortium;							
RA	Bloecker H., Boecker M., Brandt P., Newes H.W., Weil B., Amid C.,							
RA	Osanger A., Fobo G., Han M., Wiemann S.;							
RL	Submitted (SPP-2004) to the EMBL/GenBank/DBJ databases.							
DR	EMBL: AL834431; CAD39091.1.							
DR	GO: GO:0009986;							
DR	GO: GO:000624;							
DR	GO: GO:0003677;							
DR	FDR InterPro; IPR000571; Znf_CCH.							
DR	InterPro; PP00642; zf-CCCH_1.							
DR	SMART; SM00336; 2nF_C3H1; 1.							
KW	Hypothetical protein.							
FT	NON_TER	1						
SEQUENCE	728 AA;	80284 MW;	5B2E6C5682AA93A6	CRC64;				
SQ								
<b>RESULT 4</b>								
O9H5J2	ID	O9H5J2		PRT;	540	AA.		
AC	O9H5J2;							
DT	01-MAR-2001	(TREMBLrel.	16,	Created)				
DT	01-MAR-2001	(TREMBLrel.	16,	Last sequence update)				
DT	01-OCT-2002	(TREMBLrel.	22,	Last annotation update)				
DE	Hypothetical protein FLJ23389.							
OS	Homo sapiens (Human).							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
NCBI_TaxID=9606;								
RN	[1]							
RP	SEQUENCE FROM N.A.							
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemoto Y., Okamoto S.,							
RA	Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,							
RA	Tanaka T., Nakamura Y., Isogai T., Sugano S.							
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.							
DR	EMBL: AK027042; BAB15634.1.							
SQ	SEQUENCE	540 AA;	60158 MW;	8B040AFB87P17A27	CRC64;			
<b>Query Match</b>								
Best Local Similarity	59.9%	Score	3721;	DB 2;	Length	728;		
Matches	703;	Conservative	2;	Mismatches	0;	Indels	0;	
Qy								
337	VLORTGDPANTLNLRPHILELLANDDPNPDAVSPTEQLENAMYAVKTIVHGLYDFIONYS	396	Qy	652	MPYADHVYSTFSPRDMNNSPYQOPPPQYGPVPPVPSGMYAPYDVSRR1WPPMQRDDI	711		
Db	1	VLORTGDPANTLNLRPHILELLANDDPNPDAVSPTEQLENAMYAVKTIVHGLYDFIONYS	60	Db	1	MPTADHVYSTFSPRDMNNSPYQOPPPQYGPVPPVPSGMYAPYDVSRR1WPPMQRDDI	60	
Qy	397	RKGHETPQOPNSKTKTSMCRDLRQGGCPRGNTCTFHSQEELEYKIRNKKINATRT	456	Qy	712	IRSNSLPPMDVRESVTQTSLRERVNSLDGYYSVACQPSEPTVPLPREPGHLKTSC	771	
Db	61	RKGHETPQOPNSKTKTSMCRDLRQGGCPRGNTCTFHSQEELEYKIRNKKINATRT	120	Db	61	IRSNSLPPMDVRESVTQTSLRERVNSLDGYYSVACQPSEPTVPLPREPGHLKTSC	120	
Qy	457	FPLINKVGNNNTTTAGNVISVIGSTTETTGKLVPTSTNGISNAENSVQLSRSTDSTLR	516	Qy	772	BEQIRRKEDQWAQHTOKAPLYSTSPLPATQSPSPPLFSYDSEVSGTKFEBD	831	
Db	121	FPLINKVGNNNTTTAGNVISVIGSTTETTGKLVPTSTNGISNAENSVQLSRSTDSTLR	180	Db	121	BEQIRRKEDQWAQHTOKAPLYSTSPLPATQSPSPPLFSYDSEVSGTKFEBD	180	
Qy	517	ALETVKCKVGANGQAAGPSADSVTENKIGSPPKTPVSVNATAGPSNYGTTELNSVP	576	Qy	832	HLSYHSPMSCGTIGSCINAIDSEPKDVIANSNAVLMOLDSGDVKRVHLFETQRTRKED	891	
Db	181	ALETVKCKVGANGQAAGPSADSVTENKIGSPPKTPVSVNATAGPSNYGTTELNSVP	240	Db	181	HLSYHSPMSCGTIGSCINAIDSEPKDVIANSNAVLMOLDSGDVKRVHLFETQRTRKED	240	
Qy	577	QKSSPPFLTRPVYPFHSENIOYQDPRTOQPFEPYPOYDQTCGYPPPPVPGVACPYRF	636	Qy	952	SSYNEATSSAHTYERDFITDLSGRHKHSSTGSCINAIDSEPKDVIANSNAVLMOLDSGDVKRVHLFETQRTRKED	1011	
Db	301	VRSNNVPESSLPPASMPYADHYSTSFRDRMNNSPYQOPPPQYGPVPPVPSGMYAPYD	360	Db	301	SSYNEATSSAHTYERDFITDLSGRHKHSSTGSCINAIDSEPKDVIANSNAVLMOLDSGDVKRVHLFETQRTRKED	360	
Qy	697	SRR1WRPMPYQDDITRSNLLPPDMVSSVYOTSRLREYNSLDGYYSVACQPSEPTT	756	Qy	1012	QQRKNSLDEGRHTLNLSKETEFLRNGLQSDYTEDATDTKPRDIELSALTDTRDNG	1071	
Db	361	SRR1WRPMPYQDDITRSNLLPPDMVSSVYOTSRLREYNSLDGYYSVACQPSEPTT	420	Db	361	QQRKNSLDEGRHTLNLSKETEFLRNGLQSDYTEDATDTKPRDIELSALTDTRDNG	420	
Qy	757	VPLPREPCGHKLTKSCEBQIRRKEPDQWAQHTQAPLVSSTLPVATOSSTPPSLFSYD	816	Qy	1072	QSBIEELDQIQLISSONDQLINGMAVENGHVQHQKEPKKQKSLEDHVILEBK	1131	
Db	421	VPLPREPCGHKLTKSCEBQIRRKEPDQWAQHTQAPLVSSTLPVATOSSTPPSLFSYD	480	Db				

Db	4211 QSEPIEIBLDIOLGISSQNDOLLNGAVENHPVOOHPKEEPKQRQSISLGHDVILEEKQ	480
Qy	1132 TILPTISCSQSPPLVISNSACLPIITSVSGNLLTKTHMSEDKNDFLKPVANGKMNS	1191
Db	481 TILPTISCSQSPPLVISNSACLPIITSVSGNLLTKTHMSEDKNDFLKPVANGKMNS	540
RESULT 5		
Qy	Q69231 PRELIMINARY; PRT; 1136 AA.	
Db	Q69231; ID: Q69231; AC: Q69231; DT: 25-OCT-2004 (TREMBLrel. 28, Created); DT: 25-OCT-2004 (TREMBLrel. 28, Last sequence update); DT: 25-OCT-2004 (TREMBLrel. 28, Last annotation update); DB: MKIAA2025 protein (Fragment).	
GN	Name=mKIAA2025; OS: Mus musculus (Mouse); OX: Bokuryoata; Metzico; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. NCBI_TaxID=10090;	
RP	SEQUENCE FROM N.A. Okazaki N., Kikuno R.F., Ohara R., Inamoto S., Koseki H., Hiraoaka S., Seino S., Nishimura M., Kaihono K., Kitamura H., Nagase T., Ohara O., Koga H.; "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries.", DNA Res. 11:205-218(2004).	
CC	-!- SIMILARITY: Contains 1 RING-type zinc finger.	
EMBL	AK17335; BA32613.1;	
DR	GO: 0006151; GO: 0003676; GO: 0004842; GO: 0008270; GO: 0016567; InterPro: IPR000571; InterPro: IPR001841; Pfam: PF00097; PF000442; SMART: SM00184; PROSITE: PS00510; PS50089; PS50089; NON-TER	
DR	O.: Koga H.; "Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries.", DNA Res. 11:205-218(2004).	
DR	GO: 0006151; GO: 0003676; GO: 0004842; GO: 0008270; GO: 0016567; InterPro: IPR000571; InterPro: IPR001841; Zinc_finger.	
DR	Pfam: PF00097; zf-C3HC4_1; SMART: SM00184; RING_1; Zinc_finger.	
DR	PROSITE: PS00510; ZF_RING_1; PS50089; ZF_RING_2; Zinc_finger.	
DR	Metal-binding_zinc_finger.	
DR	NON-TER_1	
SQ	SEQUENCE 1136 AA; MW: 44328BA6F13CD51 CRC64;	
	Query Match 43.2%; Score 2685; DB 2; Length 1136;	
	Best Local Similarity 52.2%; Pred. No. 1.2e-125; Mismatches 295; Indels 134; Gaps 36;	
Matches 615; Conservative 135; Mismatches 295; Indels 134; Gaps 36;		
Qy	1 MVQQAOTTEFUSPCKYNEFDNVHKPISLGCSHTVKTCNLKLRKACPDQTAINTD 60	
Db	7 MVQAPONTDFUSPCKYNTQFDETIRKPIISLCGHTYCMCLINKLRKACPDQTTNTD 66	
Qy	61 IDVLPVNFALLQLVGAQYDPHQSKI-SNLGENKHVEAKCVCYDIALYLKPSCRGKVA 119	
Db	67 TELLPVNSALLQLVGAQYDPHQSKI-SNLGENKHVEAKCVCYDIALYLKPSSARGV 126	
Qy	120 -SLNQSMALSRPMQKENTLNLQVLEEGVRMARSGLCERTTELIQHNPQLS 177	
Db	127 LNSTTQSVLSRPMQKLVTLQVLEEGVRMARSGLCERTTELIQHNPQLS 186	
RESULT 6		
Q1816	PRELIMINARY; PRT; 1109 AA.	
AC	Q1816; ID: Q1816; DT: 01-MAR-2003 (TREMBLrel. 23, Created); DT: 01-MAR-2003 (TREMBLrel. 23, Last sequence update); DT: 01-OCT-2003 (TREMBLrel. 25, Last annotation update); DB: KIAA2025 protein (Fragment).	
GN	Name=KIAA2025; OS: Homo sapiens (Human).	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo sapiens	
DB	178 ANIWAAYARGCFLGPMQEALKVLAEDGSALRKVLFVQRLERFPQSKT 237	
DB	187 SNIWAAYARGCFLGPMQEALKVLAEDGSALRKVLFVQRLERFPQSKT 246	
Qy	238 SIGHVVVQLYRASCFCFKYTKRDEDSIIMOLKEEFRSYFALEHRREDAOIVTHIAEAGLISP 297	
DB	239 SIGHVVVQLYRASCFCFKYTKRDEDSIIMOLKEEFRTYALRHHSDTSQVIAEAGLIRAP 306	

OX NCBI_TAXID=9606;	Qy 727 V--YQTSRLERYNSLDGYYSVACQP---PS---BPRITVPLPRBPCGHLKTKTSCBEQIR 776
RN [1] SEQUENCE FROM N.A.	677 VPSYVPESEREYQQTESYYAPHHFQIRPXYLRSRPPSRPPPP-----HPSJBLHR 732
RC TISSUE_Brain;	
RA Nagase T., Kikuno R., Ohara O.; Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.	Qy 777 RKPDQWAQHTOKAPLYSSSTLPVATOSPTPPSPPLSVDFRADFSESSVSGTKEEDHLHSY 836
RL EMBL; AB09945; BAC23121.1;	DR 733 RRKEIMAQLEERK--VISPPPFAP-SPTPLP-PTPHPEEFLDEDKVAG-KYKGNDYSQV 786
DR HSSP; P22893; IM9O.	GO 837 SPWSCCTGSCNIAIDSEPKDVIANSNAVLMDDSDVVKRVLHETORRTEKE-BDPPI 894
DR GO; GO:000151; Ubiquitin ligase complex; IEA.	DR 787 SPWSCDTIGSYGTDAKPDVVAAGSVMENIVESKGM-RDQRDLQRRAAETSDDLI 844
DR GO; GO:0003676; F:nucleic acid binding; IEA.	GO 895 PFDGPIIISKNGAISSRSRTGXYHTDPVQATASOGSATKPISSDVYPYNAVDSRW-S 952
DR GO; GO:0008270; F:ubiquitin ligase activity; IEA.	DR 845 PFGDRPPTVSREGAISRTKRYQAGEMQANAPQAPTKSINISDPSY-GTHGGNGAS 902
DR GO; GO:0016567; P:protein ubiquitination; IEA.	DR 953 SYG--NEATSSAHYVERDRFIVTDLSGRHKR-HSSTGDLLSLELQQ----- 994
DR InterPro; IPR000571; Znf_fing.	DR 903 PYSPHQNTIPSPQCHFSERLISMSEVASHGKPLPSAREOLRLEQQLNHQISQQTQLRGL 962
DR Pfam; PF00642; Zf_CCHC_1.	GO 995 -AKNSNLLQREANALAMMQ -----KWNSLDEGRHLTNL--LSKBEIBLRNGELQSD 1043
KW Metal-binding; Zinc_finger.	DR 963 EAWSNRVLVQLREANTLAGSQSOPPPKKWGMISSEQLSLEHVOVEREIGKTRLESM 1022
PT NON_TERP 1	Db 1044 YTEDATDTKPDRIELELSALDTDE-PDGQSEPIEBILDQLGQISSQNDOLLNGMVEN 1101
SC SEQUENCE 1109 AA; 122994 MW; 1DOD6AF3E1497EC CRC64;	Db 1023 -NQCSLDMK-----SKNTNSQAOENGQEPQNQKV-----PAEDLTLTFSVDVN 1064
Query Match 41.5%; Score 2577.5; DB 2; Length 1109;	Qy 963 GHPVQHQKEPPKKQKKS 1120
Best Local Similarity 51.3%; Pred. No. 2..e-120; Indels 139; Gaps 37;	Db 1065 GSALTQENISLISLNKTSSL 1083
Matches 595; Conservative 135; Mismatches 290;	
Qy 25 VRKPISLCSHTVCKTCUNKLIRKACPPDQTAINTDIVLPNPNALQLVGAQVPDHQSI 84	
Db 1 IRKPISLCSHTVCKTCUNKLIRKACPPDQTINTDILPNTSALLQLVGAQVPDQPI 60	
Qy 85 KU-SNLGENKHYEVAKKCVEDDLALKPLSGCGKVA--SINOSALSRPMQRKLTVLNVNCQ 141	
Db 61 TLCSGVBDTQHYEARKCVEELALKPLKLLSARGVGUNLTTQSLSRPMQRKLTVLHVQ 120	
Qy 142 LVBEEGRTRAMPAARSIGERTTTELQHQNPQLSANLWAARVRCQQLGPAMQEAL 201	RESULT 7
Db 121 LVBEEGRTRAMPAARSIGERTTTELQHQNPQLSANLWAARVRCQQLGPAMQEAL 180	Q6NUC6 PRELIMINARY; PRT; 1114 AA.
Qy 202 KLVLLALEDGSALSRSRKVULFVQQLRERFPOASKTSIGHVQLLFRASCFKVTRKEDS 261	AC Q6NUC6; TRMBMLrel. 27, Created
Db 181 KLVLLALEDGSALSRSRKVULFVQQLRERFPOASKTSIGHVQLLFRASCFKVTRKEDS 240	DT 05-JUL-2004 (TRMBMLrel. 27, Last sequence update)
Qy 262 SLMQLKEEFRSYEARRHDQAQVHIAEGLRISPECWSSLYGDLAKSHMSMOSITIDLKL 321	DT 05-JUL-2004 (TRMBMLrel. 27, Last annotation update)
Db 241 SLMQLKEEFRSYEARRHDQSIVQIAEGLRISPECWSSLYGDLAKSHMSMOSITIDLKL 300	DE NGC81061 protein.
Qy 322 QSPESPAASVOELTIVLORTGDPANLNRPHIELLANDIDPFDAYSPTWOLENAMYAV 381	GN Name=NGC81061.
Db 301 QTPSPASQVOELTIVLORTGDPANLNRPHIELLANDIDPFDAYSPTWOLENAMYAV 360	OS Xenopus laevis (African clawed frog).
Qy 382 KTVVHGLDFTQNSRKHEPOPQNSKTMIRQGGCPRTGTCPTPAHSQBLE 441	OC Bokaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Db 361 RTVHGLDFTQNSRKHEPOPQNSKTMIRQGGCPRTGTCPTPAHSQBLE 420	OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
Qy 442 KYVLRNKNINATYRTPELINKVSNNTTTAGNVISVTGSTETTGKIVPS-TNGISNAE 500	OC Xenopodinae; Xenopus.
Db 421 KFQRMQKRL--VPRRPASLQGTLQRTGDPANLNRPHIELLANDIDPFDAYSPTWOLENAMYAV 477	OC NCBI_TAXID=8355;
Qy 501 NSVSQLISRSSTDSTLRALETVKVKGKADQQPPQHSSKTYKTMCRDKMQRGCCPRAASCFTAHSQBLE 560	RP SEQUENCE FROM N.A.
Db 478 NTYVQLIQRGTDP--SYDSSLKPGK1-----DHLSSAPSPDPLLESVPKS 522	RC TISSUE-Clavate; PMID=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Qy 561 TSAGPSNNTELANSVPQGSSPFLTRPVY-----YPPHSENIQYFQDPR-TQI 606	RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Altschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Diatchenko L., Matsunaga K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B., Brownstein M.J., Usdin T.B., Yoshiuki S., Carninci P., Prange C., Rah S.S., Loqueland N.A., Peters G.J., Abra罕son R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Munny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahay J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Blitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schneich A., Schein J.E., Marras M.A.; "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences";
Db 523 ISALPVNP---HSIPRPGPADLPPMPTKPLQMVPRSSQYPAQQTIV-YQDPRGAAP 577	RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
Qy 607 PFEVPOQYQGTYPPPTVPGAVCPVRFVRNNVPVSSLPASMPYADHYSTFSPRDR 666	RL RN [2].
Db 578 PFBAPYQGMYTTPP-----QCVSRFVR---PPSAPAPPFLDHYPPYL-QER 625	RP SEQUENCE FROM N.A.
Qy 667 MNNSPYQPPPQPOXPYGPVPPVPSGMYAPYDSDRIWRPPMYYQRDIIIRSNSLPPMDVNHSS 726	
Db 626 VVNSQYGTQPOQ---YPSI---YPSI-YPSI-YPSI-YPSI-YPSI-YPSI-PIEIPAA 676	



Pahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimeswood J., Schmutz J.R.M., Butterfield Y.S., Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.B., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [2]

SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA Strausberg R.; RLD to the EMBL/GenBank/DDBJ databases.

CC "- SIMILARITY: Contains 1 RING-type zinc finger.

DR EMBL; BC044642.1; -.

DR GO; GO:000151; C:ubiquitin ligase complex; IEA.

DR GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:000842; F:ubiquitin-protein ligase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0016567; F:protein ubiquitination; IEA.

DR InterPro; IPR000571; Znf\_CCH.

DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00097; zf-C3HC4\_1.

DR SMART; SM00184; ZF\_CCHC\_1.

DR PROSITE; PS00518; ZF\_RING\_1; UNKNOWN\_1.

DR PROSITE; PS00089; ZF\_RING\_2; 1.

KW Hypothetical protein\_Metal-binding\_Zinc\_Zinc-finger.

SEQUENCE 506 AA; 56853 MW; F2915DA0B8D0AC98 CRC64;

Query Match 36.9% Score 2293.5; DB 2; Length 506; Best Local Similarity 95.5%; Pred. No. 1.4e-106; Matches 448; Conservative 3; Mismatches 13; Indels 5; Gaps 1;

Qy 1 MPVQAQKTEFLSPCIVNEPDENVHKPISLGCSHTYTKTCNLKLRKACPPDOTAINTD 60 Db 1 MPVQAQKTEFLSPCIVNEPDENVHKPISLGCSHTYTKTCNLKLRKACPPDOTAINTD 60 Qy 61 IDVLPVNPFALLQLVGAQPDHQSKLUSNLLGENRYEAKKCYVEKDLYLKPLSGGGKVAS 120 Db 61 IDVLPVNPFALLQLVGAQPDHQSKLUSNLLGENKHYEAKKCYVEKDLYLKPLSGGGKVAS 120 Qy 121 LNOASLSPMQRKLVTLNQCOLVVEGRVRAARSIGERTVTELLQHNPQQLSANL 180 Db 121 LNOASLSPMQRKLVTLNQCOLVVEGRVRAARSIGERTVTELLQHNPQQLSANL 180 Qy 161 WAAVRARGCQFLGPAMQBEALKVLLAEDGALSRSRVLTVQRLPRPQQASKTSIG 240 Db 161 WAAVRARGCQFLGPAMQBEALKVLLAEDGALSRSRVLTVQRLPRPQQASKTSIG 240 Qy 241 HVVQLLYRASCFCVKTRBDSSLMQLKEFRSYEARREHDQAQVHIAEAGLRISEPW 300 Db 241 HVVQLLYRASCFCVKTRBDSSLMQLKEFRSYEARREHDQAQVHIAEAGLRISEPW 300 Qy 301 SSILYGDIAHKSHMQS1IDKLOSPESFKSVOELTIVLQRTDPANUNRLRPHLELANI 360 Db 301 SSILYGDIAHKSHMQS1IDKLOSPESFKSVOELTIVLQRTDPANUNRLRPHLELANI 360 Qy 361 DPNDPDAVSPTEOLEMANAVKTVHGLVDPIONYSRKGHEPQPQNSKTKTSMCRDLR 420 Db 361 DPNDPDAVSPTEOLEMANAVKTVHGLVDPIONYSRKGHEPQPQNSKTKTSMCRDLR 420 Qy 421 QQGCGCPRGPNCTPAHSQOBLELEYKRLRNPKINATVRTPFLNKVGVNNTV 469 Db 421 QQGCGCPRGPNCTPAHSQOBLELEYKRLRNPKINATVRTPFLNKVGVNNTV 469 Qy 421 QQGCGCPRGPNCTPAHSQOBLEKCNPRLGHLCCC---LSSYASLGTV 464 Db 421 QQGCGCPRGPNCTPAHSQOBLEKCNPRLGHLCCC---LSSYASLGTV 464

AC Q9HBBD1; DT 01-MAR-2001 (TREMBLrel. 16, Created) DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

AC Q9VVA48; DT 01-MAY-2000 (TREMBLrel. 13, Created)

AC Q9VVA48; DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

AC Q9VVA48; DE CG16807-PA.

GN ORFNames=CG16807; OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Hydroleidae; Drosophilidae; Drosophila.

OX NCBI\_TaxId=7227; [1]

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20517938; PubMed=10938276; DOI=10.1101/jbc.M004461200;

RA Siess D.C.; Veder C.T.; Merkens L.S.; Tanaka T.; Freed A.C., McCoy S.J.; Heinrich M.C.; Deffebach M.E.; Bennett R.M.; Hefnerider S.H.;

RA "A human gene coding for a membrane-associated nucleic acid-binding protein." J. Biol. Chem. 275:33655-33662 (2000).

DR EMBL; AF253304; AAG0433.1.-.

DR GO; GO:0009986; C:cell surface; IDA.

DR GO; GO:0005624; C:membrane fraction; IDA.

DR GO; GO:0003677; F:DNA binding; IDA.

FT NON\_TER 1

SQ SEQUENCE 419 AA; 46285 MW; 2C55A12C7F56DC CRC64;

Query Match 34.8%; Score 2165; DB 2; Length 419; Best Local Similarity 100.0%; Pred. No. 2.7e-100; Mismatches 0; Indels 0; Gaps 0; Matches 419; Conservative 0; MleMatches 0; DR 1 EQIRKEDQWQYHQTQAPLVSSSTLPVATQSPTPSPFLSYDFRADFSESYSGTKPEBEDH 832

Qy 773 EQIRKEDQWQYHQTQAPLVSSSTLPVATQSPTPSPFLSYDFRADFSESYSGTKPEBEDH 60

Db 1 EQIRKEDQWQYHQTQAPLVSSSTLPVATQSPTPSPFLSYDFRADFSESYSGTKPEBEDH 60

Qy 833 LSHYSPMCGTGSICNAIDEPKVDTANSVAVLMLDSDGDKYKRTHFETQRTEEDP 892

Db 61 LSHYSPMCGTGSICNAIDEPKVDTANSVAVLMLDSDGDKYKRTHFETQRTEEDP 120

Qy 893 IIPPSDGPIISKWGAISRSSRTGYHTDPVQATASQGSATKPTISDVYPTVNAVDSRWS 952

Db 121 IIPPSDGPIISKWGAISRSSRTGYHTDPVQATASQGSATKPTISDVYPTVNAVDSRWS 180

Qy 953 SYGNATSSAHYVERDRPIVTDSLGRKHSSRTGDLISLEQOKNSNLLQREALALAMQ 1012

Db 181 SYGNATSSAHYVERDRPIVTDSLGRKHSSRTGDLISLEQOKNSNLLQREALALAMQ 240

Qy 1013 QKWNLSDEGRHLTLNLSKEBLRNGLSODYTEDATDTRDTEIELSALDTDEDGQ 1.072

Db 241 QKWNLSDEGRHLTLNLSKEBLRNGLSODYTEDATDTRDTEIELSALDTDEDGQ 300

Qy 1073 SEPPIEFLDIQGISSONDOLINGAVENGHPVQOHOKEPPKKQKOSLGEDHVLBFOKT 1132

Db 301 SEPPIEFLDIQGISSONDOLINGAVENGHPVQOHOKEPPKKQKOSLGEDHVLBFOKT 360

RESULT 10

Q9VVA48 ID Q9VVA48 PRELIMINARY; PRT; 819 AA.

AC Q9VVA48; DT 01-MAY-2000 (TREMBLrel. 13, Created)

AC Q9VVA48; DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE CG16807-PA.

GN ORFNames=CG16807; OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Hydroleidae; Drosophilidae; Drosophila.

OX NCBI\_TaxId=7227; [1]

RN [1]

SEQUENCE FROM N.A. MEDLINE=22426065; PubMed=12537573;

RX Adam M.D., Celinkiner S.E., Holt R.A., Evans C.A., Gocayne J.D., Adams M.D., Celinkiner S.E., Holt R.A., Evans C.A., Gocayne J.D., George R.A., Scheer S.E., Li P.W., Hoskins R.A., Gallo R.F., Amanatides P.G., Scheer S.E., Richards S.E., Asburner S.N., Sutton G.G., Wormer J.R., Yandell M.D., Zhang Q., Chen L.X., Sutton G.G., Wormer J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H., Blazej R.G., Champé M., Pfeiffer B.D., Brandon R.C., Rogers Y.H., Blazej R.G., Champé M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Gabor G.L., Abril J.P., Agbaria A., An H.J., Andreevs Pfrankoch C., Baldwin D., Ballieu R.M., Babu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandarkar S., Borkovska D., Botchan M.R., Bouck J., Brockstein P., Brottier P., Burtis K.C., Buzdam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cowley S., Dahlke C., Davenport L.B., Davies P., dePablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., Rosler C., Gabrielsen A.B., Garg N.S., Gelbart K., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Hernandez J.R., Hernández J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Rei M.H., Ikegami C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Lai Z., Lasko P., Lei Y., Levitt A.A., Li J., Li Z., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milikhina N.V., Moriarty C., Morris J., Mosherfi A., Mount S.M., Moy M., Murphy B., Murphy J., Nelson D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M., Peinert K., Remington K., Saunders R.D., Schaefer F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Swirski R., Tectoc C., Turner R., Venter J., Wang A.H., Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J., Yeh R.P., Zaveri J.S., Woodage M., Woodage M., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C., RT "The genome sequence of *Drosophila melanogaster*," Science 287:2185-2195 (2000). [2]

RN SEQUENCE FROM N.A. MEDLINE=22426065; PubMed=12537573;

RX Celinkiner S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champé M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Patel S.M., Park S., Pfeiffer B.D., Richards S., Soderren B.J., Svirskas R., Tabor P.E., Wan K., Stagleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.; RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila melanogaster* euchromatic genome sequence," Genome Biol. 3:RESEARCH0079 (2002). [3]

RN SEQUENCE FROM N.A. MEDLINE=22426070; PubMed=12537573;

RX Kaminker J.S., Kronmiller B., Carlson J., Svirska R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celiker S.E., Lewis S.E., Rubin G.M., RT "The transposable elements of the *Drosophila melanogaster* euchromatin: a genomic perspective," Genome Biol. 3:RESEARCH0084 (2002). [4]

RN SEQUENCE FROM N.A. MEDLINE=22426069; PubMed=12537572;

RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E., Smith C.D., Tupy J.L., Whetstone E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celiker S.B., de Grey A.D., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.; RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review,"

RL Genome Biol. 3:RESEARCH0083 (2002). [5]

RN SEQUENCE FROM N.A.

RX FlyBase; Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases. [6]

RN SEQUENCE FROM N.A.

RX FlyBase; Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases. [7]

RN -!- SIMILARITY: Contains 1 RING-type zinc finger.

RX EMBL; AE003522; AA094744; CG16807; FBgn0036621; CG16807; DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.

RX FlyBase; FBgn0036621; CG16807; DR GO; GO:0003676; F:nucleic acid binding; IEA.

RX InterPro; IPR000571; Zinc\_finger. DR InterPro; IPR001841; Zinc\_finger. DR Pfam; PF00697; zf-C3HC4; 1. DR Pfam; PF00642; zf-C3HC4; 1.

RX SMART; SM00184; RING; 1. DR PROSITE; PS00518; ZP\_RING\_1; 1. DR PROSITE; PS50089; ZP\_RING\_2; 1. KW Metal-binding Zinc\_Zinc-finger.

RX SEQUENCE 819 AA; 90376 MW; 33D429076EDE7B2E CRC64;

Query Match 24.7%; Score 1538; DB 2; Length 819; Best Local Similarity 40.5%; Pred. No. 1..5e-68; Matches 386; Conservative 98; Mismatches 226; Indels 242; Gaps 29;

QY 1 MPVQAAQWTEFLSCPICYNERDENVKRPIUSGCSHTVKTCICLKHKAFCPDQTAINTD 60 DR 1 MPVQAAQWTEFLSCPICYNERDENVKRPIUSGCSHTVKTCICLKHKAFCPDQTAINTD 60

QY 61 IDVLPYTFNALLQV-----GAQVPDHOSIKLNSLGENRHVEAKKCVEDLALYLKPL 112 DR 61 IDVLPYTFNALLQV-----GAQVPDHOSIKLNSLGENRHVEAKKCVEDLALYLKPL 112

QY 62 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 292 DR 62 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 292

QY 63 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293 DR 63 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293

QY 64 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293 DR 64 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293

QY 65 LRISSPQWSLLYGLDALKHKSIMQSIDKLQSPESFAKSQVQELTIVLQRTGPANLNRLP 352 DR 65 LRISSPQWSLLYGLDALKHKSIMQSIDKLQSPESFAKSQVQELTIVLQRTGPANLNRLP 352

QY 66 LNLNG---NGNLLTRPMLRKLVTLVNCOLMEEGVRVLAARSIGERTVTELLHQHN 172 DR 66 LNLNG---NGNLLTRPMLRKLVTLVNCOLMEEGVRVLAARSIGERTVTELLHQHN 172

QY 67 PQQLSANLWAAVRARGQFQFLPAMQEBALKVLLALEDGSLASRKYTVLFLVQRLPFRP 232 DR 67 PQQLSANLWAAVRARGQFQFLPAMQEBALKVLLALEDGSLASRKYTVLFLVQRLPFRP 232

QY 68 PQLLSNLWAAVRTRGCQFLGPAMQEBVLKVLALLEGSLASRKYTVLFLVQRLPFP 233 DR 68 PQLLSNLWAAVRTRGCQFLGPAMQEBVLKVLALLEGSLASRKYTVLFLVQRLPFP 233

QY 69 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 292 DR 69 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 292

QY 70 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293 DR 70 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293

QY 71 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293 DR 71 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293

QY 72 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293 DR 72 QASKTSIGHVQLLYRASCFCVKTQRDDSLMLQKBEFRSYEARLREHDAQIVHATMEAG 293

QY 73 PQQLSANLWAAVRARGQFQFLPAMQEBALKVLLALEDGSLASRKYTVLFLVQRLPFRP 232 DR 73 PQQLSANLWAAVRARGQFQFLPAMQEBALKVLLALEDGSLASRKYTVLFLVQRLPFRP 232

QY 74 PQLLSNLWAAVRTRGCQFLGPAMQEBVLKVLALLEGSLASRKYTVLFLVQRLPFP 233 DR 74 PQLLSNLWAAVRTRGCQFLGPAMQEBVLKVLALLEGSLASRKYTVLFLVQRLPFP 233

QY 75 QPNISKYTKTSMCRDLRQGGCPRGTNTCFHSQEELFKYURKINATVRTPLLNKVG 464 DR 75 QPNISKYTKTSMCRDLRQGGCPRGTNTCFHSQEELFKYURKINATVRTPLLNKVG 464

QY 76 412 TNSNPYK1SLCRDLANRVCPRGSSCTPAHSQEEVERYRARNR-----455 DR 76 412 TNSNPYK1SLCRDLANRVCPRGSSCTPAHSQEEVERYRARNR-----455

QY 77 465 VNNTVTTAGNVTSGSTETGKIVPSTNGTSNAENNSVSQIISRSSTDSTLRALETYKV 524 DR 77 465 VNNTVTTAGNVTSGSTETGKIVPSTNGTSNAENNSVSQIISRSSTDSTLRALETYKV 524

QY 78 456 -----455 DR 78 456 -----455

QY 79 525 GRVGANGQAAGPSADSVTENKIGSPPKTPVSNVAATSAGPSNNGTTELNSVP-QKSSPP 582 DR 79 525 GRVGANGQAAGPSADSVTENKIGSPPKTPVSNVAATSAGPSNNGTTELNSVP-QKSSPP 582

QY 80 583 LTRPVYPHPHSNIQYFQDPTQIPPEVPQYPTQGYXPPIPDTVPGVAPCVPRFVRSNNV 642 DR 80 583 LTRPVYPHPHSNIQYFQDPTQIPPEVPQYPTQGYXPPIPDTVPGVAPCVPRFVRSNNV 642

Db	488 LGNMPPMLPSP-MHMGSPR-----	-GYLDP-----	-GGGL 522	
Qy	643 PESSLPPAS---MP---YADHYSTF---	-SPRDRMNSSPYQQPPPOPGPVPPVPSGMAYPV 694		
Db	523 PPSHSPITRLIVSPRYDSRFSGFGGTPR-----	-IPSPEXYQANVP-----	566	
Qy	695 YDSSRIIWPPMQRDDIIRNSLPPM--DYMHSVSYQTSLRERTYNSLDDOYSSACQPP- 750			
Db	567 --TQRNANPFSVNSNLHKGMLPASGGDVYFHLLA-----	-NPWEQAYLAQQQHPP 614		
Qy	751 -----SEPRTTVPLPREPCGHKLTKSCEQ--IRRKEQHTQKAPLYS-STLEVATQS 803			
Db	615 QHPQQQQPSKPNNSRPLSILPATADSFFEKKPNSVSDLDRLPENVDAVPLFRS 674			RESULT 12
Qy	804 PTPPSPLEFSYDFRADFSEVSIGTKFEBDHLSHYSPPSCGTGSCINAIDSBPKDVTANSN 863			Q9NNW99 PRELIMINARY; PRT; 198 AA.
Db	675 -----NNNNNNNNNSNNNNNNNGSLLFWN-NNTGKDSANFVR-----DSTLDDD 720			AC Q9NNW99; ID Q9NNW99; DT 01-OCT-2000 (TREMBLrel. 15, Created)
Qy	864 AVLMDLDSGDVKKRVLFETORRTRKEEDP1PFSDCPIISKWGA.SRSRSSTG 915			DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
Db	721 ASTFDVPTGSSMLSIY-----GPICPK-----SSTRG 747			DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
				DB Hypothetical protein FLJ0713.
				OS Homo sapiens (Human).
				OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
				OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
				OC NCB_ TaxID=9606;
				OX NCBI_TaxID=9606;
				[1]
				RN SEQUENCE FROM N.A.
				RP RA Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K.,
				RA Hirao M., Ohmori Y., Ota T., Suzuki Y., Obayashi M., Nishi T.,
				RA Shibahara T., Tanaka Y., Nakamura Y., Isogai T., Sugano S.;
				RL Submitted (FEB-2000) to the ENBL/GenBank/DDBJ databases.
				DR EMBL; AK00070; BAA91340.1; -
				SEQUENCE 198 AA; 22234 MW; E7778660861F1960 CRC64;
				Query Match 14.3%; Score 889; DB 2; Length 198;
				Best Local Similarity 98.9%; Pred. No. 5.3e-37;
				Matches 173; Conservative 2; Mismatches 0; Indels 0; Gaps 0
				OX NCB_ TaxID=180454;
				RN
				SEQUENCE FROM N.A.
				RP STRAIN=PEST;
				RC Anopheles Genome Sequencing Consortium;
				RL Submitted (APR-2003) to the ENBL/GenBank/DDBJ databases.
				CC - CAUTION: The sequence shown here is derived from an
				CC ENB/GenBank/DDBJ whole genome shotgun (WGS) entry which is
				CC preliminary data.
				EMBL; AA0000887; AA0000000000679;
				GO; GO:0003676; F:nucleic acid binding; IEA.
				Inter-Pro; IPR004070; EAA04070; EAA00571; Znf CCH.
				Pfam; PF00642; zf-CCCH; I.
				DR NON_TER 1 1
				FT NON_TER 398 398 AA; 43739 MW; F1BCAFD08815A328 CRC64;
				RESULT 13
				045962 PRELIMINARY; PRT; 1014 AA.
				ID 045962; 091898; AC 045962; 091898;
				DT 01-JUN-1998 (TREMBLrel. 06, Created)
				DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
				DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
				DE Hypothetical protein M142.6.
				GN GN_PNames=M142.6;
				OS Caenorhabditis elegans.
				OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Caenorhabditis.
				OC Peleoderaiae; Peloderinae.
				OX NCBI_TaxID=6239;
				RN
				SEQUENCE FROM N.A.
				RP STRAIN=Bries1 N2;
				RX MEDLINE=99069613; PubMed=9951916;
				RA none;
				RT RT "Genome sequence of the nematode C. elegans: A platform for
				RT investigating biology";
				RT Science 282:2012-2018(1998).
				RL S21
				RP SEQUENCE FROM N.A.

RC	STRAIN=Bristol N2;	Qy	650 ASMPYADHYSTESPRDRMNSSP-----YQPQQPYGPVPPVSSNYAPYDSSRIWRPP 704
RA	Submitted C.A.;	Db	540 APMMQATEVL----ADGOMVNQGPQRVVIMQSPTHLPGPVNMVPPQSMTPVGGP 598
RL	Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.	Qy	705 MYQRDDIIRNSNLPMDVMHSSV-YQTSLRERYNSLDGYSV--ACQPPSEPRTRY 757
RN	[3]	Db	599 M-----GPQGPMTPSIVQVPNTMWTATSPGTVIYASPPQCPPTI 643
RC	SEQUENCE FROM N.A.		
RC	STRAIN=Bristol N2;		
RA	MckMurray A.A.;		
RA	Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; 273428; CAB16481.2; JOINED.		
DR	EMBL; 299276; CA97810.2; JOINED.		
DR	EMBL; 299276; CA97810.2; JOINED.		
DR	PIR; T23764; T23764.		
DR	IntAct; O45962; WormBase; WBGene; 0010923; M142.6.		
DR	WormPep; M142.6; CE34061.		
DR	GO; GO:000151; C:ubiquitin ligase complex; IEA.		
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.		
DR	GO; GO:0008270; F:zinc ion binding; IEA.		
DR	GO; GO:0016567; B:protein ubiquitination; IEA.		
DR	InterPro; IPR001841; Znf ring.		
DR	PROSITE; PS00518; ZF RING_1; UNKNOWN_1.		
DR	PROSITE; PS00089; ZF RING_2; UNKNOWN_1.		
KW	Hypothetical protein.		
SEQ	SEQUENCE 1014 AA; 112311 MW; 7222E03B2DDE4008 CRC64;		
Qy	Query Match 13.6%; Score 843; DB 2; Length 1014;		
Qy	Best Local Similarity 30.6%; Pred. No. 1.1e-33;		
Matches	238; Conservative 99; Mismatches 278; Indels 162; Gaps 23;		
Qy	4 QAAQWTFPLSCPCYCNEPDENVEKPIKSISHTVCKTCLNKHR--KACPFPQTANTDI 61		
Db	6 QGGQWQEVLCCSNRHNHE-TPLVPSLICGHVTCRKAEGKDENOTCPCHDWKTTSP 64		
Qy	62 DVLPLVNPAILQLYGAQVDPDHOSIKLNSLGENKHVEAKCVDALTYLKPLSGGKGVASL 121		
Db	65 SEVNNVALLSVI---FPRKQCMTLGSVASEAKRVDSLQSI-JAKPREADSERG-GTV 119		
Qy	122 NQSLSRPQMQRQLVLTUJNCQVLFEEGRVTRAMPAARSLSGERTYTELILQHQNQQQLSANLW 181		
Db	120 SREBISRTLQRKVTLALLCYQWREVDGRIKTLKMRGQSIIQSMTHVSSLQW 179		
Qy	182 AA VRARGQFLGPMQBALKVLLAEDGALSRSKYLFLVQRLERPPRQA SKTISGH 241		
Db	180 SAVRARGQFLGPMQBALKVLLAEDGALSRSKYLFLVQRLERPPRQA SKTISGH 239		
Qy	242 WQQLYRASCPCFKVTKRDKDSIMOLKEFRSFEYALRREHHDAQVHIAEAGRISPEQWS 301		
Db	240 WQQLYRASCFCNVLKRDKESSNLQLEFERTESLRREHDSQVQIAFESGRIGPDDQS 299		
Qy	302 SLYGDLAHKSHMOSITDKLQSPESPAKSVOELTIVLQRTGDPAN-----LNRLRP 352		
Db	300 ALIYADQBRSHMOSITDKLQSPESPAKSVOELTIVLQRTGDPAN-----LNRLRP 356		
Qy	353 HLELLANDDPNPDAVSPTEOLENAMEVAKTVVHGLDFIONY--SRKGHETPPOPOWS 409		
Db	357 CIEFFAGIEEHEDTSRMIGDALHQIRLKKNATVRTEFLINKVGNTV 414		
Qy	410 KYKTSMCRLDROOGCPRGTNCTPAHSOBELEYRNKINATVRTEFLINKVGNTV 469		
Db	415 -----NMGGPQSG-----GAEAGRI-----430		
Qy	470 TTAGNVTSVIGSETTGTKIVPSTNGISNAENVSOLSRSTUDSTLRALEYTKVKVKGVA 529		
Db	431 -----GGLPLYSQDETERSISSTNPKNHSNPQ-----461		
Qy	530 NGQNAAGSADSUTENKGSPPKTPVNSTAATSAGPSMVGTELNSVQKSSPPLTRVPPV 589		
Db	462 -----TPPKOPRQKRYQMGIPPPM-----YSSDAPPFISHQQ 497		
Qy	590 PPHSENIVQFQDPRTQIPPFVQVQPTGYYPPPTVPPGVAPCVPRVRSNNVPESSLPP 649		
Db	498 PPP----QFPNS---QHLPQRFRGGQRGAPPBPQPP-----MPMLIGYD-----MPG 539		

RESULT 14  
 ID Q96F27 PRELIMINARY; PRT; 177 AA.  
 AC Q96F27; PubMed=12477932; DOI=10.1073/pnas.242303899;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RX MEDLINE=22388257; PubMed=22388257; DOI=10.1073/pnas.242303899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemesh C.M., Schuler G.D., Altenschul S.F., Zeeberg B., Buetow K.H., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonal M.F., Casavant T.R., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Raha S.S., Loquellano N.A., Peters G.J., Prange C., Blakesley R.W., Touchman J.W., Green E.D., Abramson R.D., Mullahy S.J., Bosak S.A., McEvany P.H., McKernan K.J., Maler J., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Heaton E., Kettman J., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Dickson M.C., Schmitz J.D., Butterfield Y.S., Krzywinski M.I., Skalska U., Schein J.E., Jones S.J., Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Uterus;  
 RA Strausberg R.; Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 19669 MW; 286EBCA8EC735581 CRC64;  
 Query Match 12.6%; Score 780; DB 2; Length 177;  
 Best Local Similarity 10.0%; Pred. No. 1.3e-31; Mismatches 0; Indel 0; Gaps 0;  
 Matches 153; Conservation 0;  
 DR EMBL; BC011688; AAH11688.2; -.  
 KW Hypothetical protein.  
 FT NON TER 1 1  
 SQ SEQUENCE 177 AA; 1966

RESULT 15	QBN9V1	PRELIMINARY:	PRT:	522 AA.	Qy	906 GAISSRSRTGYHTTDPVQATASQGSAATKPIVSVDYPPYNAVDSRW--SSYG--NEATSS 961
QD	OBN9V1;				Db	269 GAISRSRTKTYQGAGMOMAAPGAKTNSINSDSPY--GTHGGWGAESPQHNPISQ 326
AC					Qy	962 AHYVERDRFIVTDLGSHRK--HSSTGDLLSBLQQ--AKNSNLILQR 1004
DT	01-OCT-2002 (TRMBLrel. 22, Created)				Db	327 GHFSERERISMSEVAASHGKPLPSAREQRLRELQQLNHOISQQTQLRGPEAVSNRLVQR 386
DT	01-OCT-2002 (TRMBLrel. 22, Last sequence update)				Qy	1005 BANALAMQQ----KWNSLDEGRHLTLNL--JSKEIEURNGELOSDYTEDATOTKPD 1054
DT	01-OCT-2002 (TRMBLrel. 22, Last annotation update)				Db	387 BANTLAGQSQQPPPWPQMISQSLSLHQREIGKRTRBLUME-NOCSDLMK--443
DB	Homo sapiens (Human).				Qy	1055 RDIBELSLADLTDE- PDGQSEBPIEBILDOLGISSQNDDQLNGMAVENGHPVQOHQKEP 1112
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Db	444 ----SKUNTSKOAEENGQPEPONKV-----PAEDLTFTSDVPNGSALTQENISL 488
OC	Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.				Qy	1113 PKOKRSL 1120
OX					Db	489 LSNKTSLL 496
RN	[1]	SEQUENCE FROM N.A. TISSUE=Testis;			Qy	Search completed: August 4, 2005, 01:40:03 Job time : 197 sec
RP	PubMed:14702039; DOI=10.1038/ng1285;				Db	
RX	RA	Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyana T., Irie R.,			Qy	
RA	Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,				Db	
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,				Qy	
RA	Yamamoto J., Kawai Y., Isono Y., Nagashiro K., Murakami K., Yasuda T., Iwayangsi T., Wagatsuma M., Shiratori A.,				Db	
RA	Sudo H., Hosoiiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi T., Kanda K., Furuya T., Kikkawa B., Omura Y., Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki K., Ninomiyra K., Ishibashi T., Yamashita H., Murakawa K., Ono Y., Takiuchi S., Watanabe S., Yosida M., Hotuta T., Ishida S.,				Qy	
RA	Kanehori K., Takahashi-Fuji A., Hara H., Tanase T., Nomura Y., Togiya S., Koma I. F., Hara R., Takeuchi K., Arita M., Imose N., Musashino K., Yuki K., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Nakagawa S., Momiyama N., Sato N., Takami S., Terashima Y., Suzuki O., Hishigaki H., Watanabe T., Sugiyama A., Kawanami B., Yamauchi M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Ikeda Y., Kobatake N., Inagaki Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeita K., Seriba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Sato T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagae T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.; human "Complete sequencing and characterization of 21,243 full-length human cDNAs"; Nat. Genet. 36:40-45 (2004). DR ENBL; AR093501; BAC04186.1; SQ SEQUENCE 522 AA; 58253 MW; D4D774D3D00DD769 CRC64;			Db		
Qy	9.8%	Score 611; DB 2; Length 522;			Qy	Best Local Similarity 34.1%; Pred. No. 1. 6e-22;
Matches 187; Conservative 76; Mismatches 187; Indels 98; Gaps 27;					Db	50 Q---YPPP-----QCVSFVFR----PSAPEAPPVDPHYDPI-PIEIPAAVPSYVPEER 100
Qy	618 YPPPPPTVAGVAPCPVPRVSNVNPESLPPASMPYADHYSTFSPRDMNSSPYQOPPP 677				Qy	678 QPYGPVPPVPGMAYPVDSRRILWRPMYQDDIIRNSLPPMDMEISSV--YQTSLER 735
Db	2 YPPP-----QCVSFVFR----PSAPEAPPVDPHYDPI-PIEIPAAVPSYVPEER 100				Db	50 Q---YPPP-----QCVSFVFR----PSAPEAPPVDPHYDPI-PIEIPAAVPSYVPEER 100
Qy	736 YNSLDGYYSVACOP----PS---EPRTTVPLPREPCGHLKTSCEBQIRKBDQWAQHT 787				Qy	736 YNSLDGYYSVACOP----PS---EPRTTVPLPREPCGHLKTSCEBQIRKBDQWAQHT 787
Db	101 YQQIESYPVAPHPQTQPSLRLPPQP--HPSLDELHRRKEIMAQLEE 156				Db	101 YQQIESYPVAPHPQTQPSLRLPPQP--HPSLDELHRRKEIMAQLEE 156
Qy	788 QKAPLYSSSTLPVATOSPPIPSPFLSVDFSESVSGTKFEDHLSHYSPWSCTGTCSC 847				Qy	788 QKAPLYSSSTLPVATOSPPIPSPFLSVDFSESVSGTKFEDHLSHYSPWSCTGTCSC 847
Db	157 RK---VIEPPPFA-PSPLP-PITHPEFLDEBKLVAG-KYKNDYQSQSPSCDITG 210				Db	157 RK---VIEPPPFA-PSPLP-PITHPEFLDEBKLVAG-KYKNDYQSQSPSCDITG 210
Qy	848 INALDSEPKDVIANSNAVMDLDSGDFYRVLHETORTKE--EDPILIPFDGPIIISKW 905				Qy	848 INALDSEPKDVIANSNAVMDLDSGDFYRVLHETORTKE--EDPILIPFDGPIIISKW 905
Db	211 IGPDKAKPKDVAAGSVEMMNVESKGM--RDQRLDLQARRAETSDDDLIIPCDRPTVSRF 268				Db	211 IGPDKAKPKDVAAGSVEMMNVESKGM--RDQRLDLQARRAETSDDDLIIPCDRPTVSRF 268

This Page Blank (usptc)